



# **Progenesis LC-MS Fractionation User Guide**

**Analysis workflow guidelines**

**for version 4.0**



## Contents

Introduction .....	3
How to use this document .....	3
How can I analyse my own runs using LC-MS? .....	3
LC-MS Data used in this user guide.....	3
Workflow approach to LC-MS run analysis .....	4
Overview of a typical Fractionation Analysis workflow .....	5
Analysing a single fraction.....	6
Stage 1: Data import and QC review of LC-MS data set for a fraction .....	7
Stage 2: Reference Run selection.....	7
Stage 3: Alignment .....	8
Stage 4: Filtering .....	10
Stage 5: Experiment Design Setup for Analysed Runs.....	11
Stage 6: Validation, review and editing of results .....	12
Stage 7: Identify Peptides.....	13
Stage 8: Refine Identifications.....	15
Stage 9: Resolve Conflicts .....	16
Completion of Fractionation Workflow Step 1 .....	19
Fractionation Workflow Step 2 .....	20
Stage 1 Import Fractions .....	20
Stage 2 Recombine Samples.....	22
Stage 3 Experiment Design Setup .....	24
Stage 4 Protein View .....	25
Stage 5: Protein Statistics .....	27
Stage 6: Reporting.....	28
Stage 7 Saving a Multi-Fraction experiment .....	30
Appendix 1: Manual assistance of Alignment.....	32
Appendix 2: Within-subject Design.....	37

## Introduction

This user guide takes you through the processes involved in the analysis of a fractionated label free LC-MS experiment. Where the method of fractionation can be gel based on electrophoresis: 1D, 2D, Off gel etc or chromatography using an additional LC step prior to the LC-MS.

In this example, to demonstrate the processes at each stage, an experiment using samples that have been fractionated using ion exchange chromatography is described. As each fraction constitutes the full application of the label free workflow as described in the main LC-MS user guide a shortened version of these analysis steps are described here.

It starts with LC-MS data file loading then Alignment, followed by Analysis that creates a list of interesting features (peptides) which can be explored within Progenesis Stats using multivariate statistical methods, then onto Protein identity.

## How to use this document

This document is designed to be used as a guide to the processes involved in the analysis of fractionated samples. Currently a full data set is not provided as this would result in a considerable download. The initial section of the document is concerned with an abbreviated description of the main experimental workflow as applied to a single fraction (a more comprehensive description of the main analysis workflow is available in the main User guide). The second section describes the process of recombining these individual fraction experiments into a 'Multi-fraction' experiment.

## How can I analyse my own runs using LC-MS?

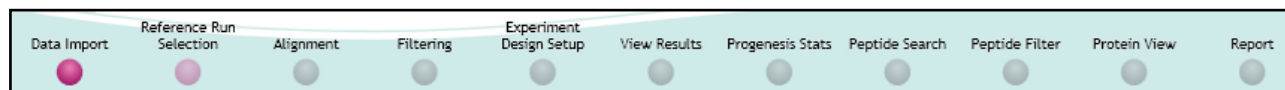
You can freely explore the quality of your LC-MS data using Data Import and then licence your own LC-MS runs using this evaluation copy of Progenesis LC-MS. Instructions on how to do this are included in a section at the end of the user guide document. Alternatively if you would like to arrange a demonstration in your own laboratory contact [support@nonlinear.com](mailto:support@nonlinear.com) and we will help you.

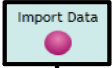

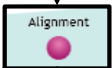



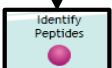
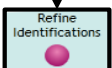


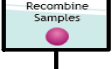
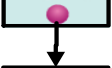


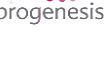
## LC-MS Data used in this user guide

NLD would like to thank Dr Robert Parker and Prof Haroun Shah at the Health Protection Agency, London, UK for providing the example data used in this user guide as well as invaluable discussion on the handling of the data.

## Workflow approach to LC-MS run analysis

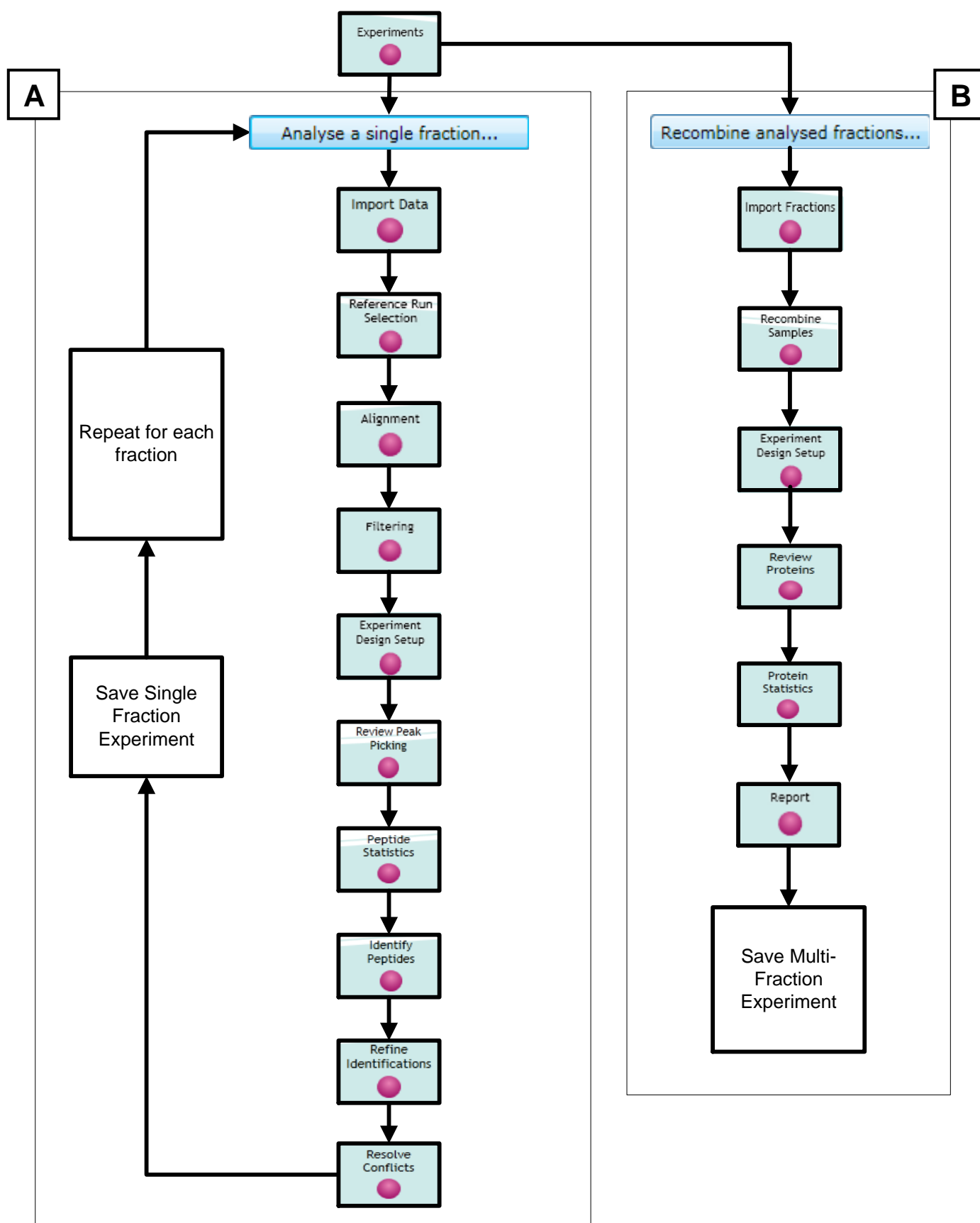
Progenesis LC-MS adopts an intuitive **Workflow** approach to performing comparative LC-MS data analysis. The following user guide describes the various stages of this workflow (see below) focusing mainly on the analysis of fractionated samples.



Stage	Description	Page
	<b>LC-MS Import Data:</b> Selection and review of data files for single fraction analysis.	7
	<b>Reference Run Selection:</b> Select run to align to.	7
	<b>Alignment:</b> automatic and manual run alignment	8
	<b>Filtering:</b> defining filters for features based on Retention Time, m/z , Charge and Number of Isotopes.	10
	<b>Experiment Design Setup:</b> defining one or more group set ups for analysed aligned runs	11
	<b>Review Peak Picking:</b> review and validate results, edit peak detection, tag groups of features and select features for further analysis	12
	<b>Identify Peptides:</b> managing export of MS/MS spectra to, and import of peptide ids from Peptide Search engines	13
	<b>Refine Identifications:</b> manage peptide ids and filters	15
	<b>Resolve Conflicts:</b> validation and resolution of peptide id conflicts for data entered from Database Search engines	16
	<b>Import fractions:</b> import multiple analysed fractionated experiments	20
	<b>Recombine samples:</b> regenerate samples from fractions	22
	<b>Experiment design Setup:</b> define original experimental design	24
	<b>Review Proteins:</b> review protein and peptide expression and identity	25
	<b>Protein Statistics:</b> perform multivariate statistics on Proteins	27
	<b>Report:</b> generate reports on proteins of interest	28

## Overview of a typical Fractionation Analysis workflow

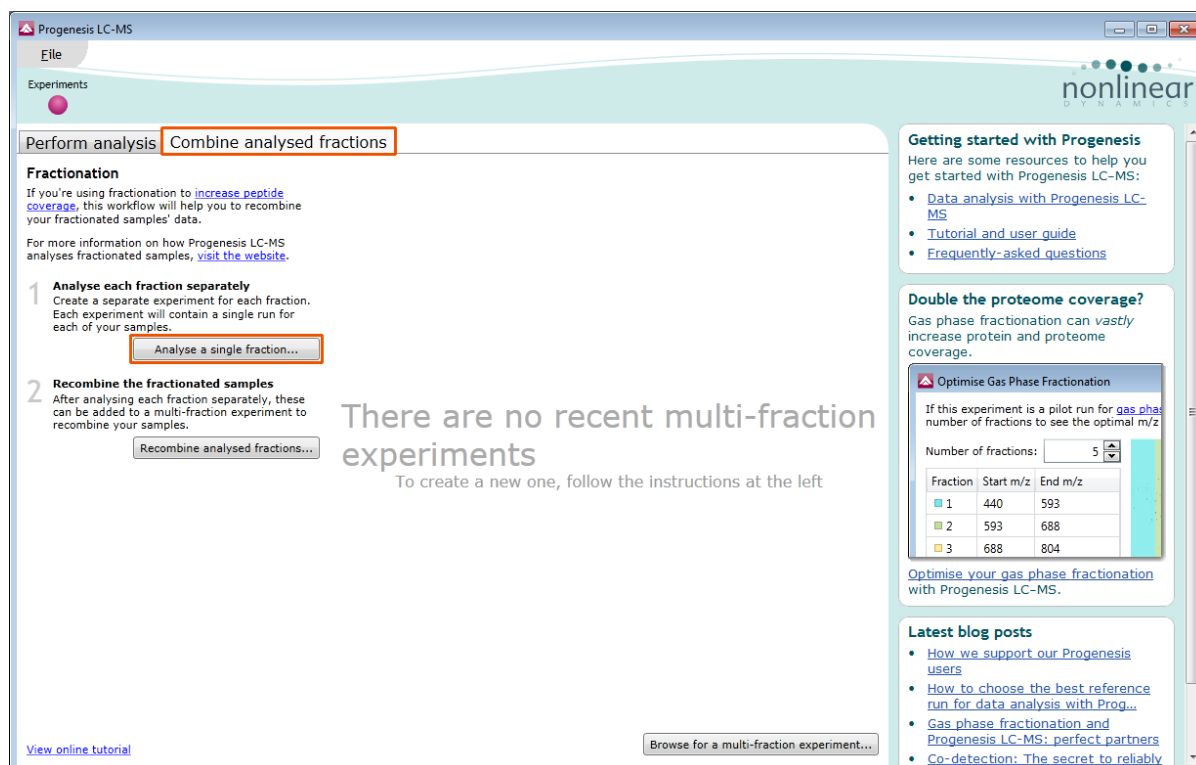
The workflow outlines the steps performed in the analysis of a typical fractionated experiment. In this example there were 6 fractions hence the main analysis workflow was performed 6 times. To analyse the data, select the **Combine analysed fractions** tab then (A) Analyse a single fraction then (B) Recombine the analysed fractions into a multi-fraction experiment.



## Analysing a single fraction

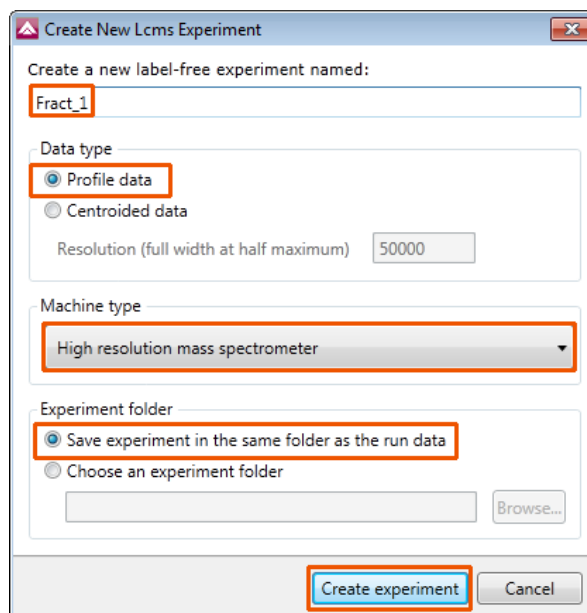
Open Progenesis LC-MS and click on the **Combine analysed fractions** tab to start the processing of Fractionated Samples.

To start the analysis of the LC-MS runs for a fraction, click on **Analyse a single fraction...**



This opens the 'Create New Experiment' dialog.

Name the fraction to be analysed (**Fract\_1**) then adjust the Data and Machine types accordingly and set the Experiment folder as required.

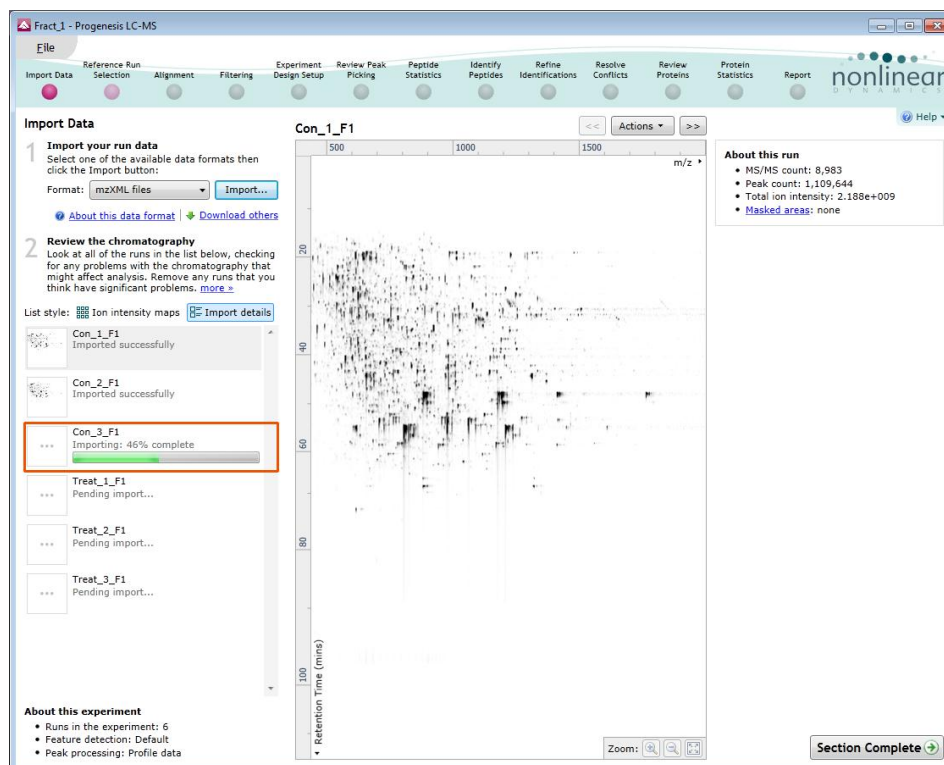


**Note:** current settings shown are the defaults

Click **Create experiment**

## Stage 1: Data import and QC review of LC-MS data set for a fraction

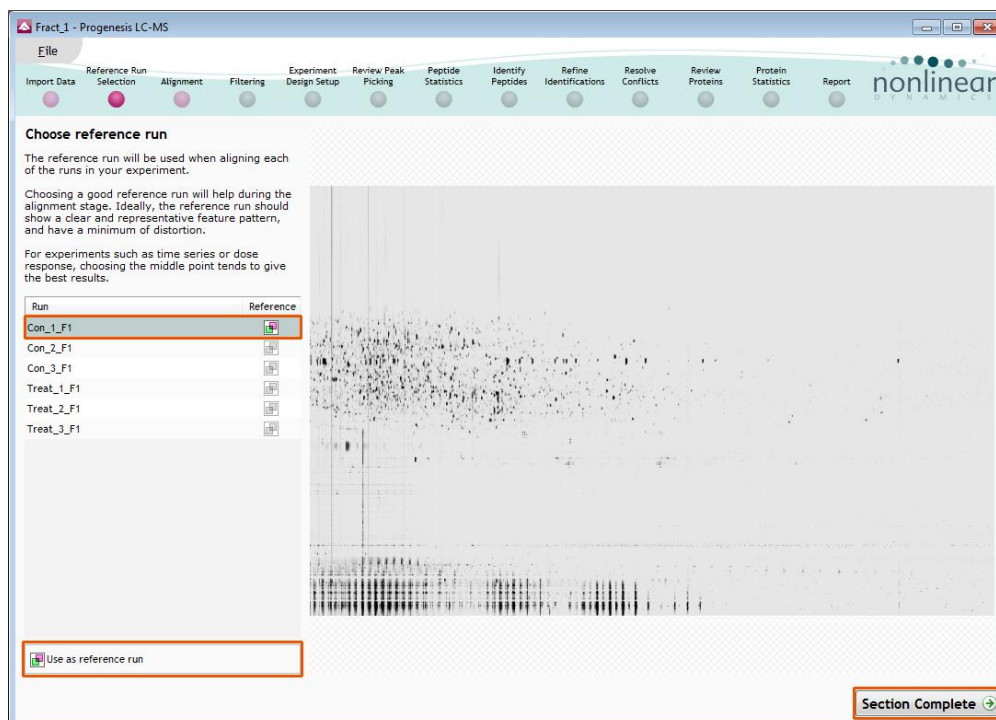
All LC-MS runs for the first Fraction are loaded into a new LC-MS experiment.



Once all the files have been imported move to the next stage in the workflow by clicking **Section Complete**.

## Stage 2: Reference Run selection

This stage in the analysis workflow allows you to review and select the most appropriate Reference LC-MS run to align all the other runs to.

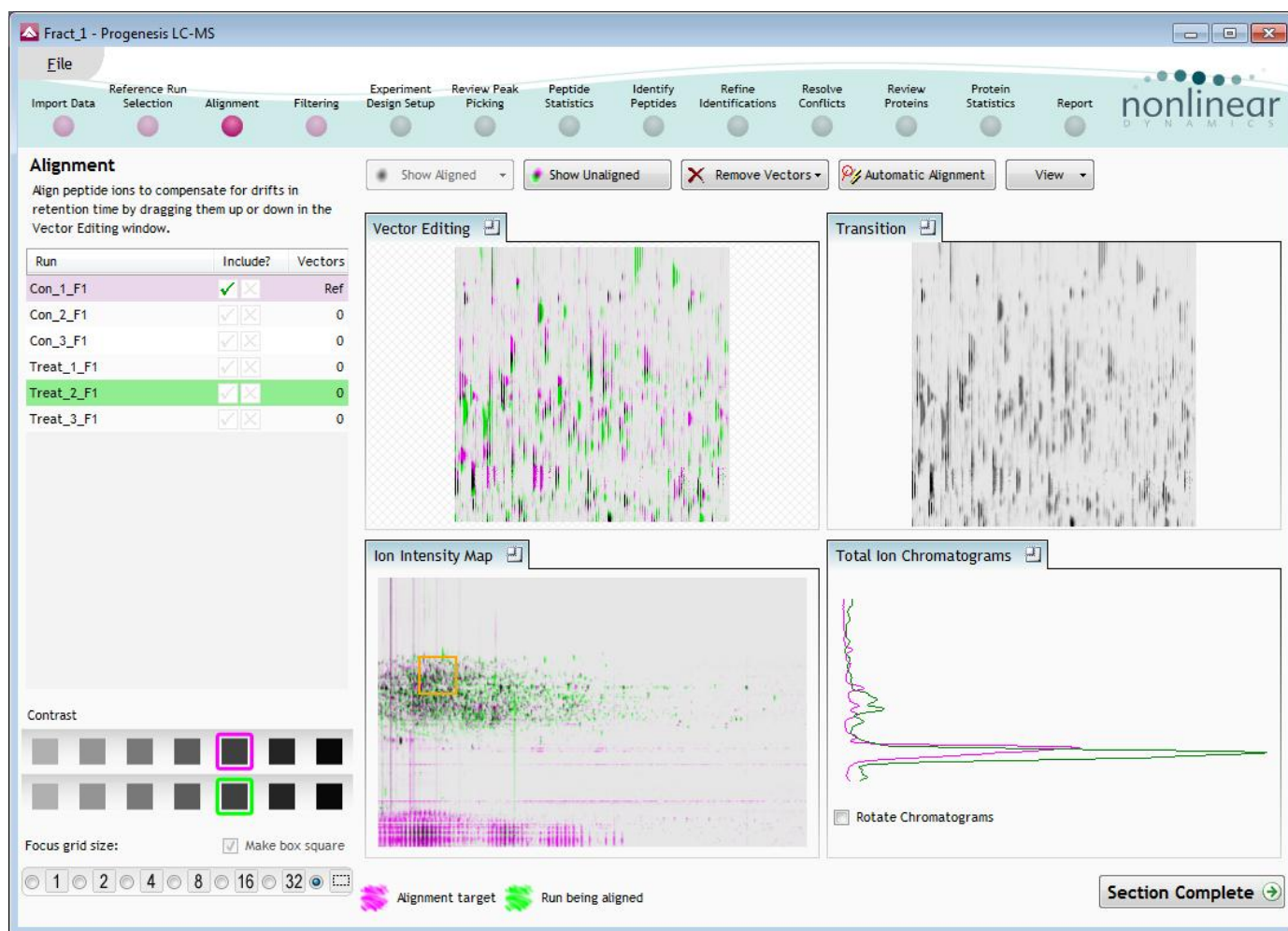


To select a Reference run either click on the run in the list and then click **Use as reference run** or double click on the run in the list.

Now move to the next stage in the workflow by clicking **Section Complete**.

## Stage 3: Alignment

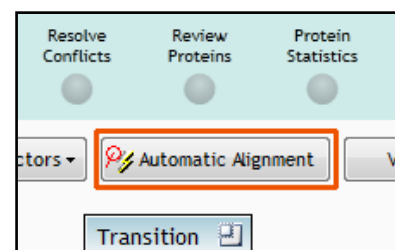
At this stage Progenesis LC-MS Alignment opens displaying your data.



### Generation of alignment vectors

The alignment of LC-MS runs is required in the LC (retention time) direction, this is key to correcting for the variable elution of peptides during the chromatographic separation.

The alignment vectors are generated automatically for all the LC-MS runs by using the 'Automatic vector wizard' accessed by clicking on **Automatic Alignment** on the top tool bar, and making sure all the runs are selected before clicking OK.





## Reviewing generation of alignment vectors

After applying **Automatic alignment** the number of vectors will be updated on the **Runs** panel and the vectors will appear (in blue) on the view.

If the alignment has worked well then in Windows A and C the grid lines should show minimal distortion, Window B (Transition) will show features pulsing slightly but not moving up and down.

**Fract\_1 - Progenesis LC-MS**

**Alignment**

Align peptide ions to compensate for drifts in retention time by dragging them up or down in the Vector Editing window.

**Runs**

Run	Include?	Vectors
Con_1_F1	<input checked="" type="checkbox"/>	Ref
Con_2_F1	<input checked="" type="checkbox"/>	157
Con_3_F1	<input checked="" type="checkbox"/>	93
Treat_1_F1	<input checked="" type="checkbox"/>	93
Treat_2_F1	<input checked="" type="checkbox"/>	102
Treat_3_F1	<input checked="" type="checkbox"/>	112

**Vector Editing** (A)

**Transition** (B)

**Ion Intensity Map** (C)

**Total Ion Chromatograms**

**Focus grid size:** 1 2 4 8 16 32 (selected)

☒ Make box square

**Section Complete**

At this point, you should check the automatically placed (blue) vectors. This will be easier with a larger grid size. Make sure the grid size is set to 4 using the '**Focus grid size**' control at the bottom left of the window.

In each square, you can, if required edit the vectors to improve the run alignment (for more detailed information on performing the alignment of your runs refer **Appendix 1** (page 32) and also to the main LC-MS analysis User guide.

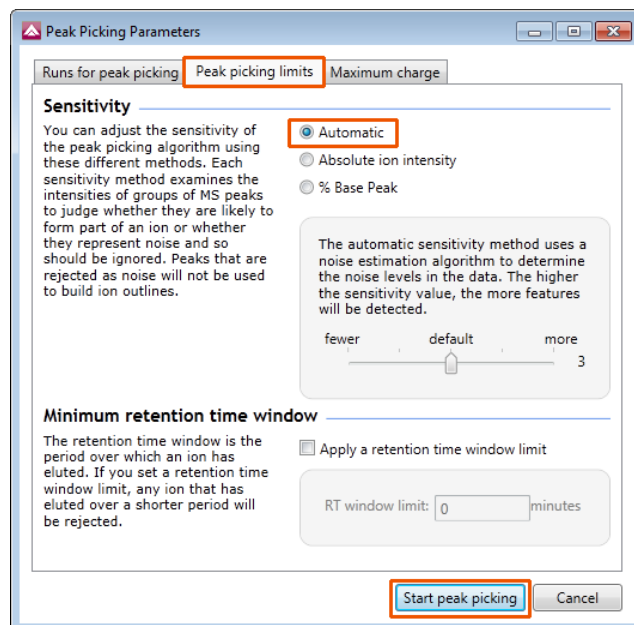
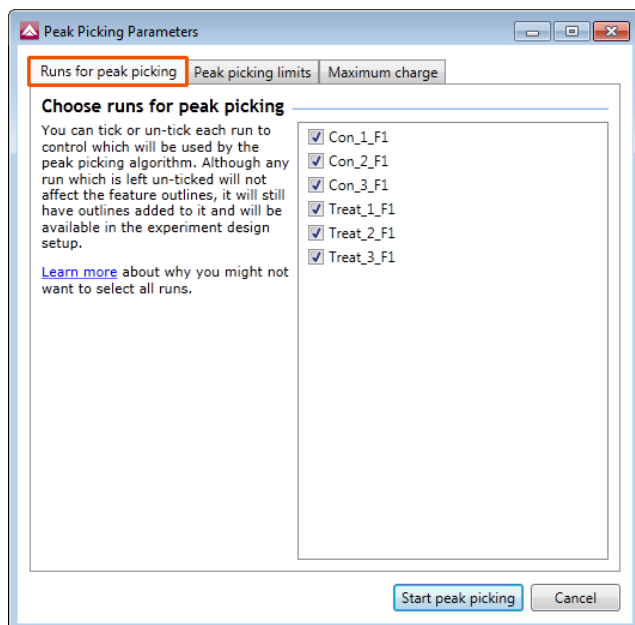
## Stage 4: Filtering

Now that you have reviewed your aligned Runs, you are ready to analyse them. Move to the **Filtering** stage, by either clicking on **Section Complete** (bottom right) or on Filtering on the workflow.



### Peak Picking Parameters

The Peak Picking Parameters dialog opens, showing all the runs in the current experiment and a tick against each run. This is the default setting, where the peak picking algorithm uses information from all of the runs to contribute to the pattern of feature outlines.



**Tip:** It may be appropriate **only** to pick peaks that are present in a limited number of your runs. In which case un-tick the runs that you do **NOT** want to contribute to the feature detection pattern. This may be important when one or more of the runs appear noisy due to non-optimal chromatography or sample handling.

**Note:** feature outlines will be added to 'un-ticked' runs; although these runs will not contribute to the peak picking pattern.

**Tip:** depending on run quality, a suggested minimum number of ticked runs should include at least one replicate of each experimental condition.

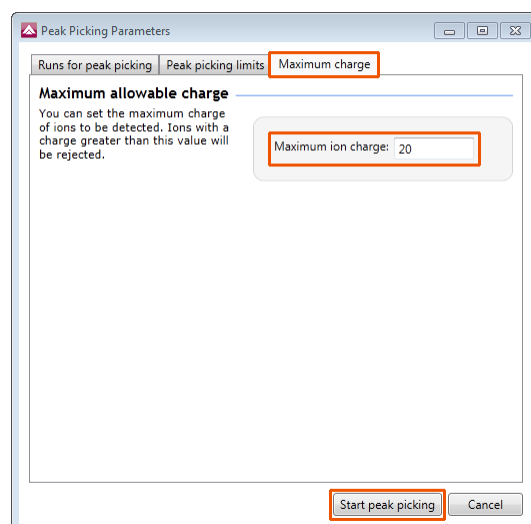
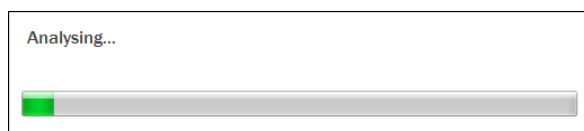
The sensitivity of the detection can be controlled by adjusting settings under the **Peak picking limits** tab.

For this example the default settings for the **Automatic** method were used.

The third tab allows you to set the **maximum charge** of the ions that will be detected. The default setting is a charge state of 20.

Press **Start Peak Picking** to start the detection process.

During the few minutes that the automatic analysis requires, a progress bar will appear telling you that it is Analysing.



More details on the management of sensitivity are available in the How to do on **Adjusting the Sensitivity of Feature detection**.

On completion of analysis the Filtering stage will open displaying the number of features.

If required you can remove features based on position, charge state, number of isotopes or combinations of these feature properties.



The example here shows the removal of features with a charge state of **less than 2 and greater than 7** and outside a defined retention time window.

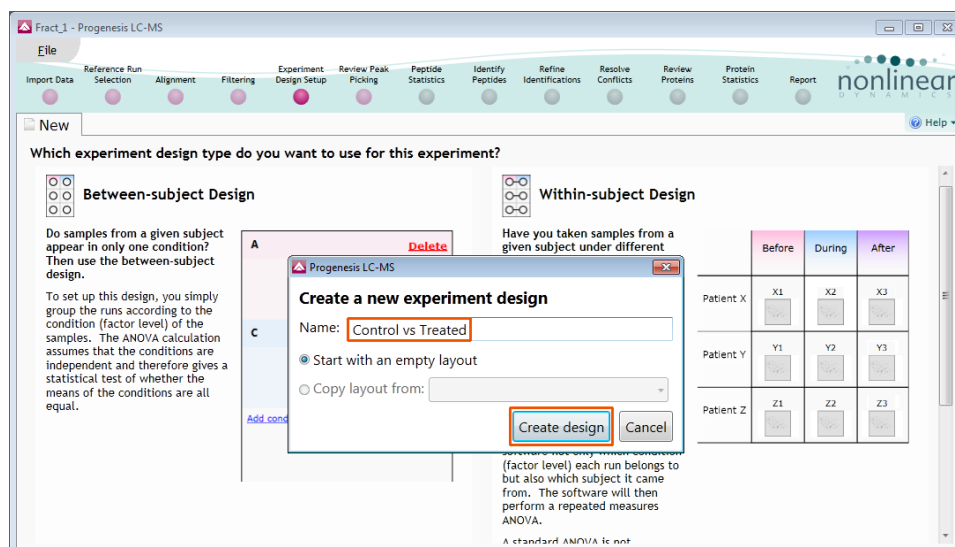
Having removed features the Normalisation will recalculate as you move to the next section.

## Stage 5: Experiment Design Setup for Analysed Runs

At this stage in the workflow you can setup one or more experimental designs for your data.

There are two basic types of experimental designs:

**Between-subject design:** where samples from any given subject appear in only one condition.

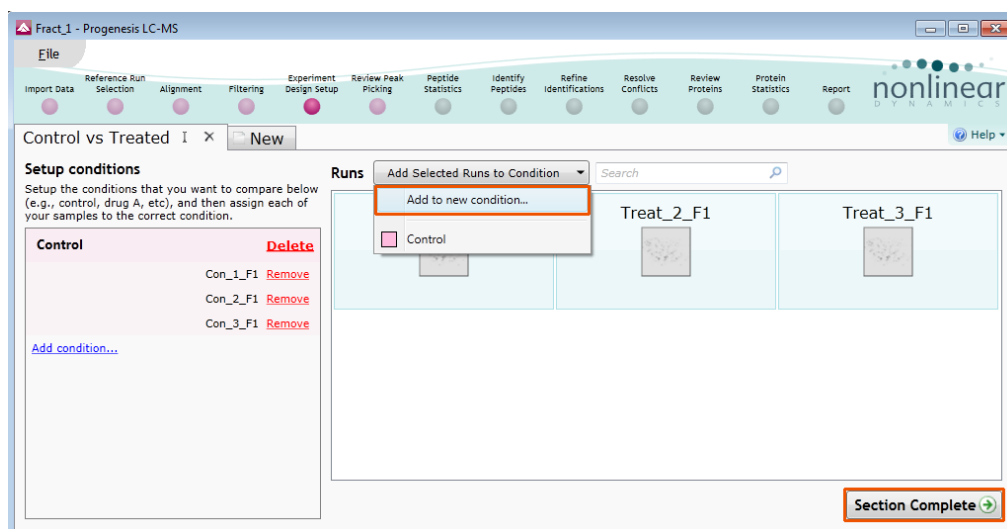


**Within-subject design:** where samples have been taken from a given subject under different conditions (Additional information on how to apply the Within-subject Design is in **Appendix 2** page 37)

This experiment contains 2 conditions: Control and Treated and uses the **Between-subject design** to group the analysed runs to reflect the Biological conditions in the original study.

Select **Between-subject** and give design an appropriate name.

Highlight the runs, to add them on to a new condition by clicking on **Add Selected Runs to Condition**

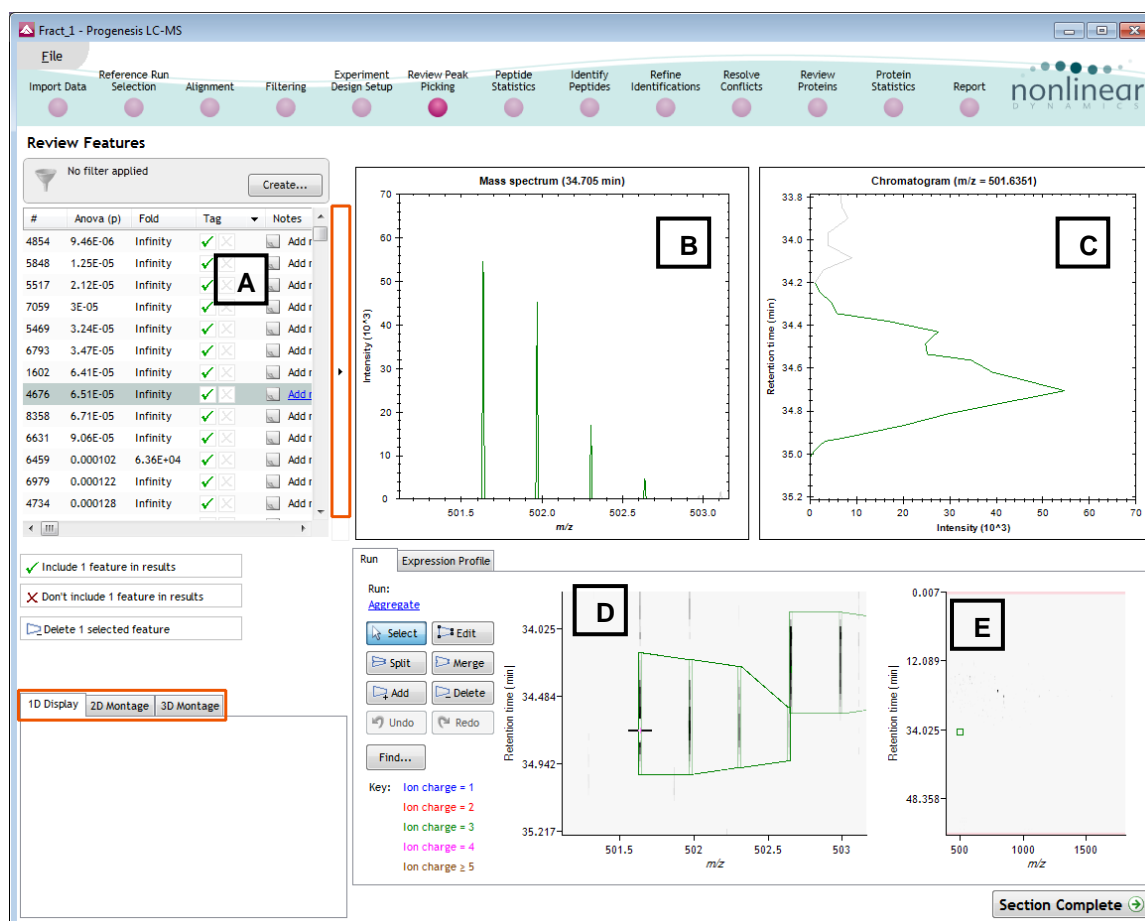


Click **Section Complete** to move to View Results.

## Stage 6: Validation, review and editing of results

The purpose of this stage in the Workflow is to review the list of features using the visual tools provided and edit features if required.

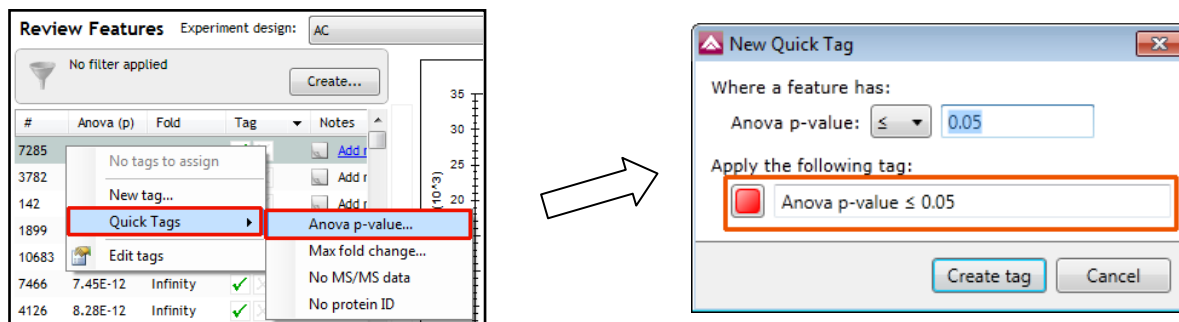
The review stage has 4 display modes: 1D, 2D, 3D and Feature Details controlled by the tabs on the bottom left of the display and the expander bar to the right of the table. Each display has multiple views to allow comparative exploration of the detected features on the aligned LC-MS runs.



Details on how to use the various views and table are described in detail in the Main analysis workflow userguide.

For the purposes of this example we require to identify all those features that demonstrate a significant Anova value ( $p \leq 0.05$ ) between the 2 conditions being studied. We will create a Tag identifying just those features.

Right click on a feature in the table and select **Quick Tags** then **Anova p-value..** set the required threshold and either accept the tag name or overwrite it.

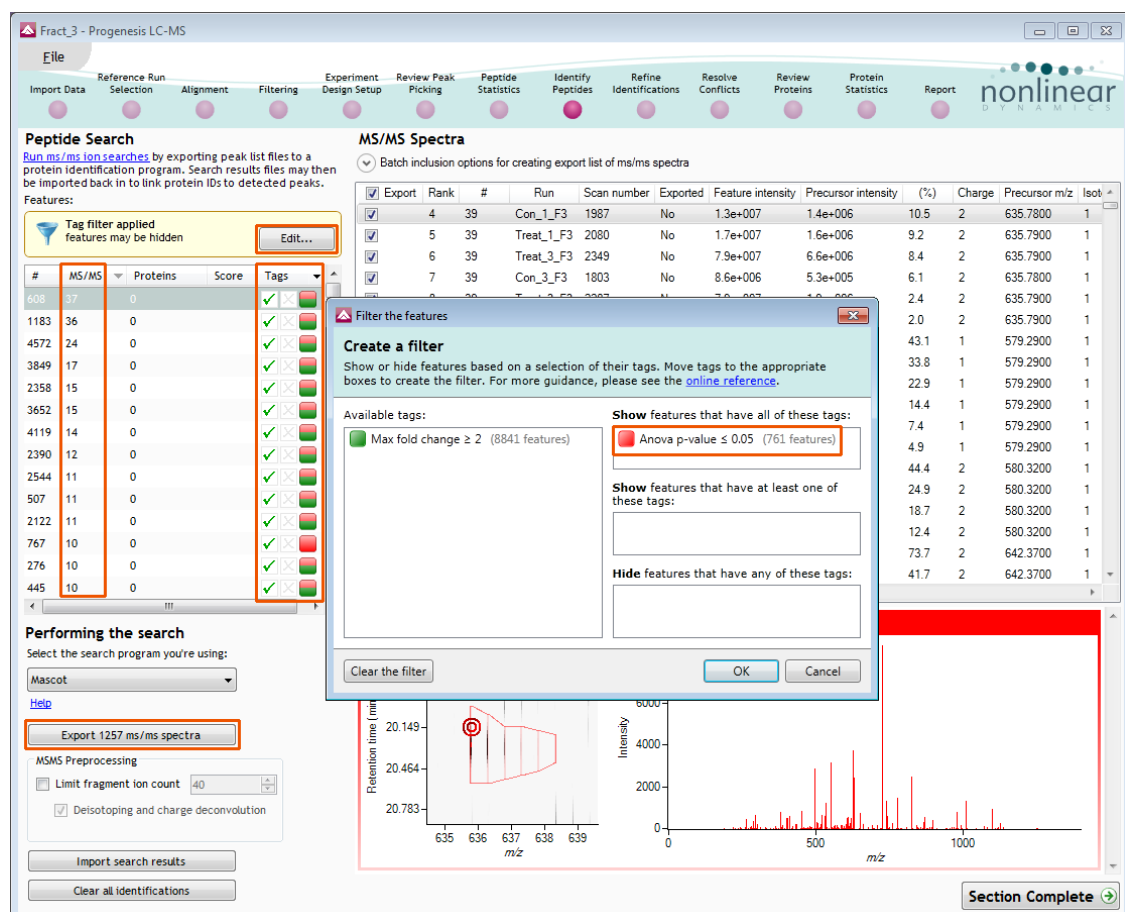


Now move to the **Peptide Search** stage in the workflow using the icon on the workflow.

## Stage 7: Identify Peptides

Progenesis LC-MS does not perform peptide identifications itself. Instead it supports identifications by allowing you to export a set of MS/MS peak lists in formats which can be used to perform peptide searches by various search engines. The resulting identifications can then be imported back into Progenesis LC-MS, using a number of different file types, and matched to your detected features.

Determining protein identification is dependant on the availability of MS/MS data for the LC-MS runs. This data may be available but limited if the LC-MS was performed in a data dependant MS/MS detection mode due to under sampling. Under these conditions MS/MS data acquisition is dependant on thresholds and parameters set prior to the acquisition of the LC-MS run.



For this example we are using LC-MS runs containing MS/MS data where the data was acquired in a data dependant mode.

The Peptide Search page shows the number MS/MS that have been matched to each feature in the Feature list (see above). MS/MS scans are matched to a feature if their precursor  $m/z$  and aligned retention time fall within the area of one of the isotopes of the feature. The MS/MS scans which are matched to the displayed features are shown in the MS/MS spectra list on the right.

The first step is to decide which MS/MS scans you wish to export to be identified. By default this is all the available spectra for the Features displayed in the Features list. Using the tag created in the previous section you can filter the table to only those showing a significant change (Anova  $p \leq 0.05$ ) between the conditions. This number of spectra to be exported is visible on the Export button.

## Performing an MS/MS Ion Search

1. Select appropriate search engine i.e. Mascot
2. Click 'Export current query set' to save search as file
3. Perform search on appropriate search engine and save results file
4. Click 'Import search results', locate results file and open
5. On importing the Search results the Features table updates to reflect the identified proteins and the relevant score for each searched feature.

**Performing the search**

Select the search program you're using:

Mascot

[Help](#)

Export 0 ms/ms spectra

MS/MS Preprocessing

☐ Limit fragment ion count 40

☒ Deisotoping and charge deconvolution

**Import search results**

Clear all identifications

Fract\_3 - Progenesis LC-MS

File Import Data Reference Run Selection Alignment Filtering Experiment Design Setup Review Peak Picking Peptide Statistics Identify Peptides **Refine Identifications** Resolve Conflicts Review Proteins Protein Statistics Report

**Peptide Search**

Run ms/ms ion searches by exporting peak list files to a protein identification program. Search results files may then be imported back in to link protein IDs to detected peaks.

Features:

Tag filter applied  
Features may be hidden

#	MS/MS	Proteins	Score	Tags	N
608	37	3 B4E2X2_H...	9.42	✓	
1183	36	3 B3K598_H...	10.5	✓	
4572	24	6 PEO1_HU...	5.35	✓	
3849	17	3 RBM5_HU...	13	✓	
2358	15	0		✓	
3652	15	3 MB81A_HU...	9.46	✓	
4119	14	4 A6NIH1_H...	6.66	✓	
2390	12	1 CTNA2_HU...	0.78	✓	
2544	11	4 B3KMC6_H...	9.2	✓	
507	11	5 B3KQ33_H...	2.3	✓	
2122	11	1 D3DU32_H...	3.64	✓	
767	10	0		✓	
276	10	4 SF3B3_HU...	53.4	✓	
445	10	5 PTAD1_HU...	87.5	✓	
3325	10	4 A2A3R6_H...	1.57	✓	

**MS/MS Spectra**

Batch inclusion options for creating export list of ms/ms spectra

Rank less than

Feature ID less than

Charge less than

Scan number less than

Exported equal to

Isotope less than

ID score less than

Feature intensity less than

Precursor intensity less than

Precursor intensity (%) less than

Run name contains

Peptide sequence contains

Protein accession contains

Protein description contains

Include in export Exclude from export Clear all filters

Export	Rank	#	Run	Scan number	Exported	Feature intensity	Precursor intensity	(%)	Charge	Precursor m/z	Isotope	Id
<input type="checkbox"/>	4	39	Con_1_F3	1987	Yes	1.3e+007	1.4e+006	10.5	2	635.7800	1	68.2
<input type="checkbox"/>	5	39	Treat_1_F3	2080	Yes	1.7e+007	1.6e+006	9.2	2	635.7900	1	87.2
<input type="checkbox"/>	6	39	Treat_3_F3	2349	Yes	7.9e+007	6.6e+006	8.4	2	635.7900	1	91.5
<input type="checkbox"/>	7	39	Con_3_F3	1803	Yes	8.6e+006	5.3e+005	6.1	2	635.7800	1	67.1

**Performing the search**

Select the search program you're using:

Mascot

[Help](#)

Export 0 ms/ms spectra

MS/MS Preprocessing

☐ Limit fragment ion count 40

☒ Deisotoping and charge deconvolution

**Import search results**

Clear all identifications

**Feature number 39, m/z 635.7849, retention time 20.269 min, charge +2**

Run: Con\_1\_F3 Scan number: 1987

Section Complete

In order to review, and refine the quality of the **Search** results click on the next stage in the workflow, **Refine Identifications**.



## Stage 8: Refine Identifications

In this example the organism under study is *Clostridium difficile*

As an **example** 'Acceptance Criteria' on which to base the sequential filtering of the Peptide results, the following thresholds will be applied:

- Remove identifications with a Score less than 30
- Remove identifications where less than 2 hits were returned
- Remove all identifications where the Protein Description **Contains** 'hypothetical'
- Remove all identifications where the Protein Description **Doesn't contain** 'Clostridium difficile'

On the Batch detection options panel, set the Score to less than 30, then **Delete matching search results**.

The screenshot displays the Progenesis LC-MS software interface. The 'Peptide Search Results' panel is active, showing 'Batch deletion options' where 'Score' is set to 'less than 30'. The 'Delete matching search results' button is highlighted. A dialog box prompts the user to confirm deleting 5463 search results. The main table lists search results with columns for #, Score, Hits, m/z, RT(mins), Charge, Mass, Mass error, Sequence, Accession, and Modifications. The bottom status bar indicates '7422 search results, 5463 matching batch delete options.'

**Note:** the search results matching the filter criteria turn pink and the number of search results being deleted is displayed

Now **Clear all filters** and then apply the next filters as described above.

To validate the Peptide search results at the protein level select the next stage in the workflow by clicking on **Protein View**.

## Stage 9: Resolve Conflicts

The Protein View combines the quantitative LC-MS data with the qualitative MS/MS results at the protein level, highlighting proteins of interest between experimental groups. This stage allows you to examine the behaviour of the identified peptides and resolve any conflicts for the various peptide assignments at the protein level.

The Resolve Conflicts view provides a number of interrelated graphical and tabular views to assist you in the validation of the peptides that have been assigned to proteins and also to review the relevance of the data returned from the search.

Open the Resolve Conflicts and order the data in the Proteins table (A) on the basis of **Conflicts**.

**Note:** the look of the tables (with regards to ordering) in the following section may vary slightly.

**Proteins**

No filter applied Create...

Accession	Peptides	Conflicts	Score	Anova (p)*
gi 5668937	12	14	1.41E+03	2.03E-05
gi 126697810	9	14	1.13E+03	4.55E-05
gi 209571234	24	13	2.4E+03	4.05E-05
gi 260682215	23	13	2.03E+03	6.28E-05
gi 126698450	12	9	1.21E+03	4.05E-06
gi 126700407	9	9	1.04E+03	0.945
gi 255656776	9	9	925	0.00888
gi 255654924	7	9	645	8.66E-06

**Peptides of gi|5668937**

No filter applied Create...

#	Score	Hits	Mass	Mass error (p...	RT (mins)	Charge	Tags	Abundance	Conflicts	Peptide
446	93.3	10	1676.838	1.2	34.5	2	✓	7.47E+05	1	IRDT
3335	84.7	4	1423.65	0.405	22.5	2	✓	7.39E+04	1	DTD
147	101	10	1230.609	0.44	22.7	2	✓	3.07E+06	1	AAD
166	125	10	2317.115	0.168	38.7	2	✓	5.55E+06	1	LES
179	60.9	9	2317.115	0.201	38.7	3	✓	3.09E+06	1	LES
238	107	10	1716.857	0.429	30.4	2	✓	1.73E+06	1	VNT
564	51.2	4	1716.858	0.394	30.4	3	✓	3.57E+05	1	VNT
283	49.5	10	1676.838	1.32	34.5	3	✓	7.59E+05	1	TRD

Protein: gi|5668937 flagellin [Clostridium difficile]  
 Protein: gi|126697810 flagellin subunit [Clostridium difficile 630]

**Peptide Views** **Protein Resolution**

**Conflicting proteins for feature 446**

Accession	Peptides	Conflicts	Protein Score	Peptide
gi 5668937	12	14	1.41E+03	93.3
gi 126697810	9	14	1.13E+03	93.3

**Peptides of gi|126697810**

#	Score	Hits	Mass	Mass error (p...	RT (mins)	Charge	Tags	Abundance	Conflicts	Peptide
147	101	10	1230.609	0.44	22.7	2	✓	3.07E+06	1	AADD
166	125	10	2317.115	0.168	38.7	2	✓	5.55E+06	1	LEST
179	60.9	9	2317.115	0.201	38.7	3	✓	3.09E+06	1	LEST
238	107	10	1716.857	0.429	30.4	2	✓	1.73E+06	1	VNTN
564	51.2	4	1716.858	0.394	30.4	3	✓	3.57E+05	1	VNTN
283	49.5	10	1676.838	1.32	34.5	3	✓	7.59E+05	1	IRDT
446	93.3	10	1676.838	1.2	34.5	2	✓	7.47E+05	1	IRDT
431	49.6	8	1692.835	0.206	20.6	3	✓	5.56E+05	1	IRDT
789	103	10	1692.833	1.09	20.6	2	✓	4.62E+05	1	IRDT
525	104	10	1700.863	0.139	36	2	✓	5.46E+05	1	VNTN
1166	42.3	7	1700.863	0.0934	36	3	✓	1.03E+05	1	VNTN
1413	99	6	1407.657	1	38.3	2	✓	1.29E+05	1	DTDV

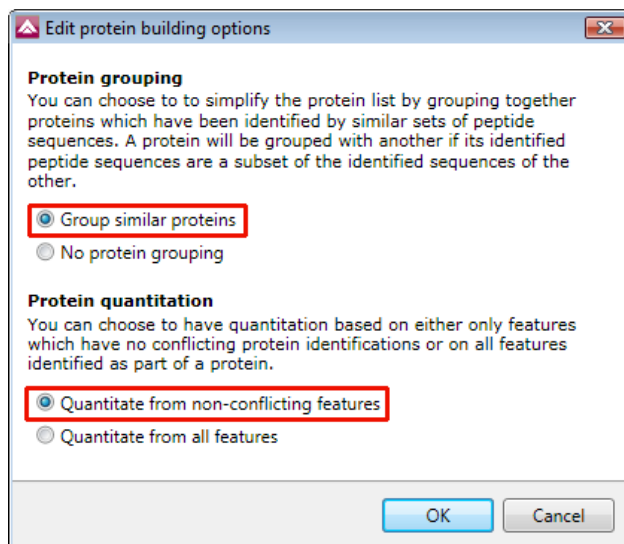
Protein options... Section Complete

In the above example the conflict would be resolved in favour of the protein with 12 peptides as the flagellin subunit does not contain any unique peptides as compared to flagellin.

**Note:** the number of conflicts you have to resolve will depend on the scope and stringency of the filters you apply at the **Peptide Search** stage,

**Note:** with the default **Protein options** for protein grouping and Protein quantitation are set as shown.





Having performed the conflict resolution with **Group similar proteins** and **Quantitate from non-conflicting features** now switch off the protein grouping.

As grouping is switched off the grouped proteins appear with conflicts to the other group members and the number of unique peptides that are used for quantitation appear in brackets after the peptide number.

**Proteins** No filter applied Create...

Accession	Peptides	Conflicts	Score	Anova (p)*
gi 5668937	12 (3)	14	1.41E+03	4.15E-06
gi 126697810	9 (0)	14	1.13E+03	---
gi 209571234	24 (12)	13	2.4E+03	2.51E-07
gi 260682215	23 (11)	13	2.03E+03	4.08E-05
gi 126698450	12 (5)	9	1.21E+03	1.28E-06
gi 126700407	9 (2)	9	1.04E+03	0.000764

**Peptides of gi|5668937** No filter applied Create...

#	Score	Hits	Mass	Mass error (p...)	RT (mins)	Charge	Tags	Abundance	Conflicts	Peptide Seq
446	93.3	10	1676.838	1.2	34.5	2	✓	7.47E+05	1	IRDT
3335	84.7	4	1423.65	0.405	22.5	2	✓	7.39E+04	1	DTT
147	101	10	1230.609	0.44	22.7	2	✓	3.07E+06	1	AAD
166	125	10	2317.115	0.168	38.7	2	✓	5.55E+06	1	LES
179	60.9	9	2317.115	0.201	38.7	3	✓	3.09E+06	1	LES
238	107	10	1716.857	0.429	30.4	2	✓	1.73E+06	1	VNT

↑ Protein: gi|5668937 flagellin [Clostridium difficile]  
↓ Protein: gi|126697810 flagellin subunit [Clostridium difficile 630]

**Peptide Views** Protein Resolution

**Conflicting proteins for feature 446**

Accession	Peptides	Conflicts	Protein Score	Peptide
gi 5668937	12 (3)	14	1.41E+03	✓ 93.3
gi 126697810	9 (0)	14	1.13E+03	✓ 93.3

**Peptides of gi|126697810**

#	Score	Hits	Mass	Mass error (p...)	RT (mins)	Charge	Tags	Abundance	Conflicts	Peptide Seq
147	101	10	1230.609	0.44	22.7	2	✓	3.07E+06	1	AAD
166	125	10	2317.115	0.168	38.7	2	✓	5.55E+06	1	LES
179	60.9	9	2317.115	0.201	38.7	3	✓	3.09E+06	1	LES
238	107	10	1716.857	0.429	30.4	2	✓	1.73E+06	1	VNT
564	51.2	4	1716.858	0.394	30.4	3	✓	3.57E+05	1	VNT
283	49.5	10	1676.838	1.32	34.5	3	✓	7.59E+05	1	IRDT
446	93.3	10	1676.838	1.2	34.5	2	✓	7.47E+05	1	IRDT
431	49.6	8	1692.835	0.206	20.6	3	✓	5.56E+05	1	IRDT
789	103	10	1692.833	1.09	20.6	2	✓	4.62E+05	1	IRDT
525	104	10	1700.863	0.139	36	2	✓	5.46E+05	1	VNT

Protein options... Section Complete

With protein grouping switched on protein groups and the additional members are indicated by a bracketed number located after the Accession number. Taking **flagellin** as an example, when the cursor is held over the accession number the group members appear in a tool tip.

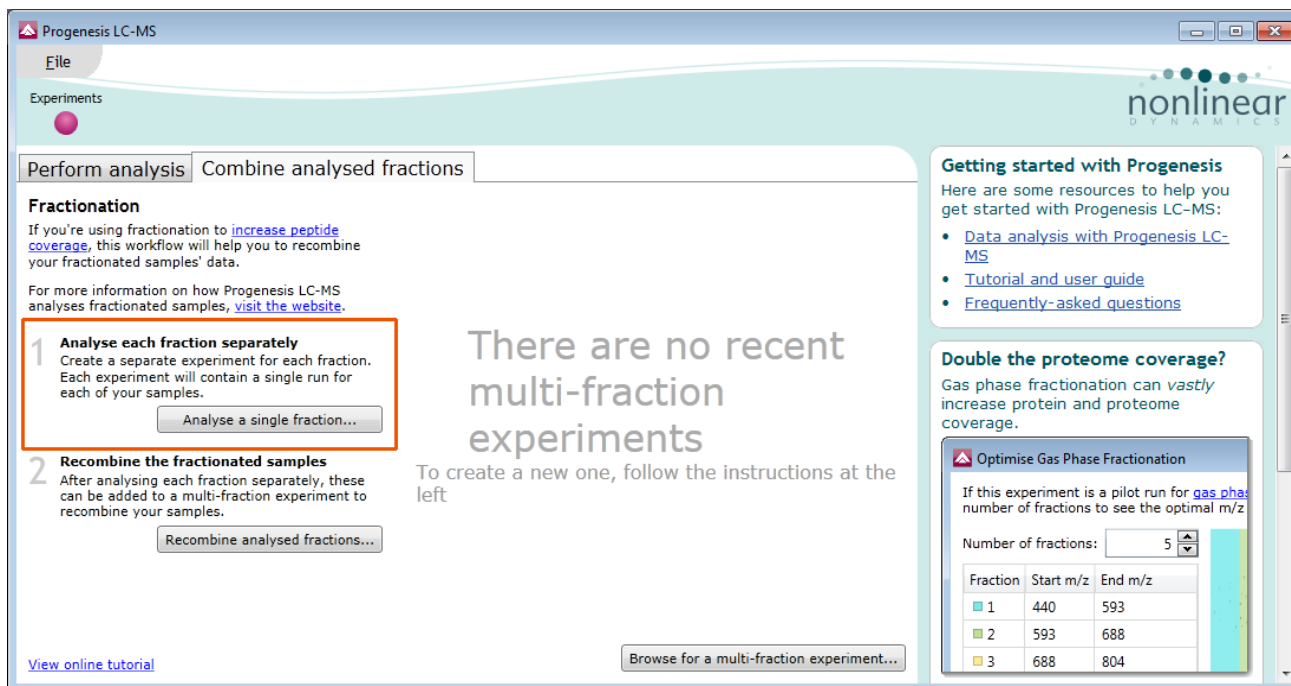
The screenshot displays the Progenesis LC-MS Fractionation software interface. The 'Proteins' table on the left lists various proteins, with 'gi|5668937 (+1)' highlighted. A tooltip for this protein shows its details: 'gi|5668937 - flagellin [Clostridium difficile]' and 'gi|126697810 - flagellin subunit [Clostridium difficile 630]'. The 'Peptides of gi|5668937' table on the right shows a list of peptides with their scores, hits, and masses. Below these tables, the 'Conflicting proteins' and 'Peptides of conflicting protein' sections are empty. At the bottom, there is a 'Protein options...' button and a 'Section Complete' button.

**Note:** the flagellin subunit has **no unique** peptides (brackets after the peptides field in the Proteins table as shown above) as they are all present in flagellin protein hence the reason for grouping. As a result all the conflicts are internal to the group.

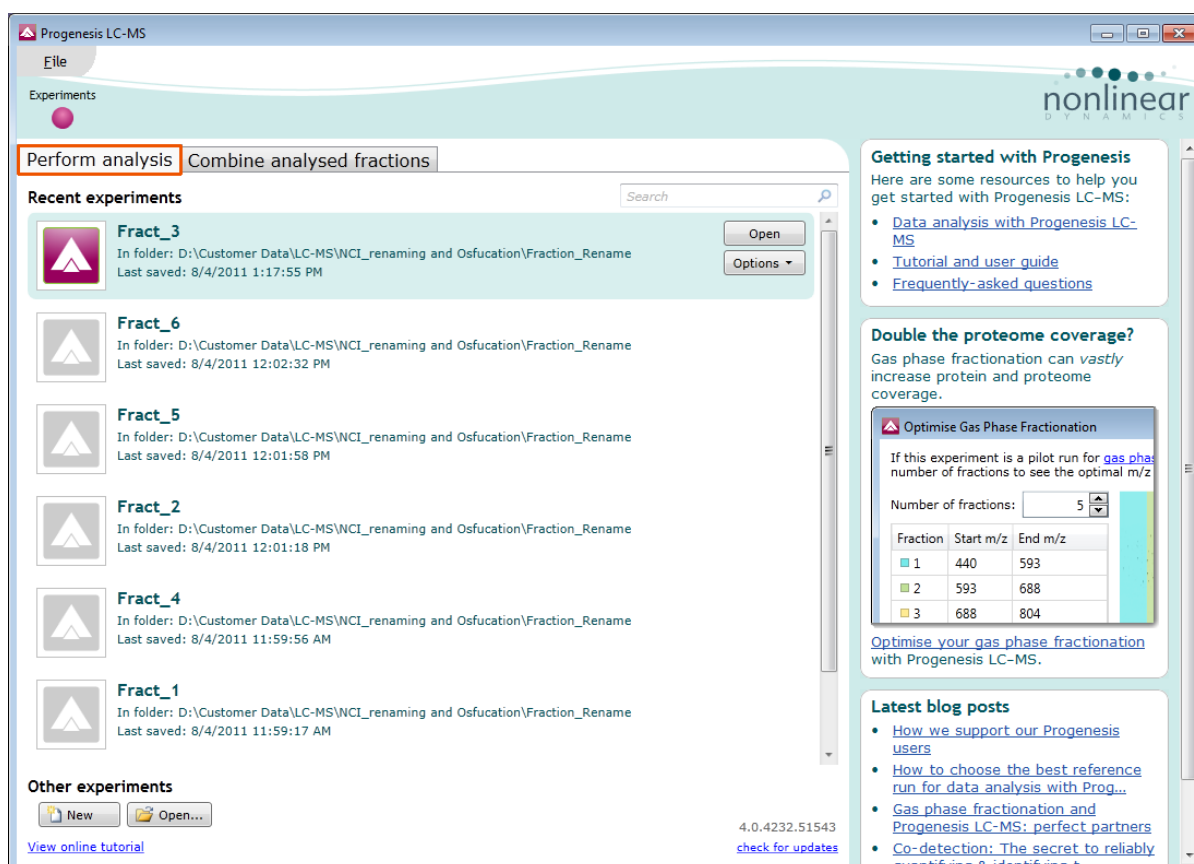
Before recombining the fractions make sure all the single fraction experiments are complete.

## Completion of Fractionation Workflow Step 1

The first stage in the analysis of a Fractionated experiment is completed when you have analysed all the single-fraction experiments, in this example there are 6 fractions, therefore 6 experiments. For each experiment this includes the identification of proteins and resolution of any peptide conflicts as described in the previous sections.



The six experiments will appear in the Perform Analysis tab and can be accessed individually.

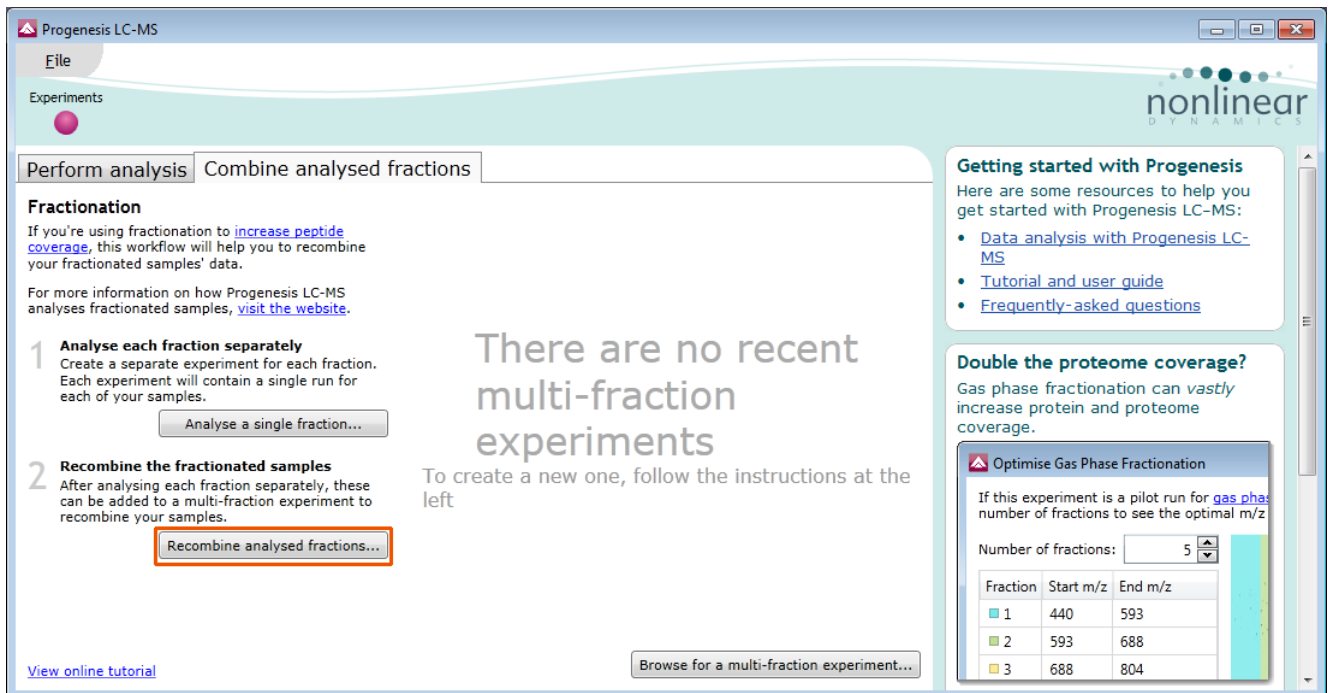


The next section describes the workflow involved in the 'Combining' of these single fraction experiments to generate a 'multi-fraction' experiment.

## Fractionation Workflow Step 2

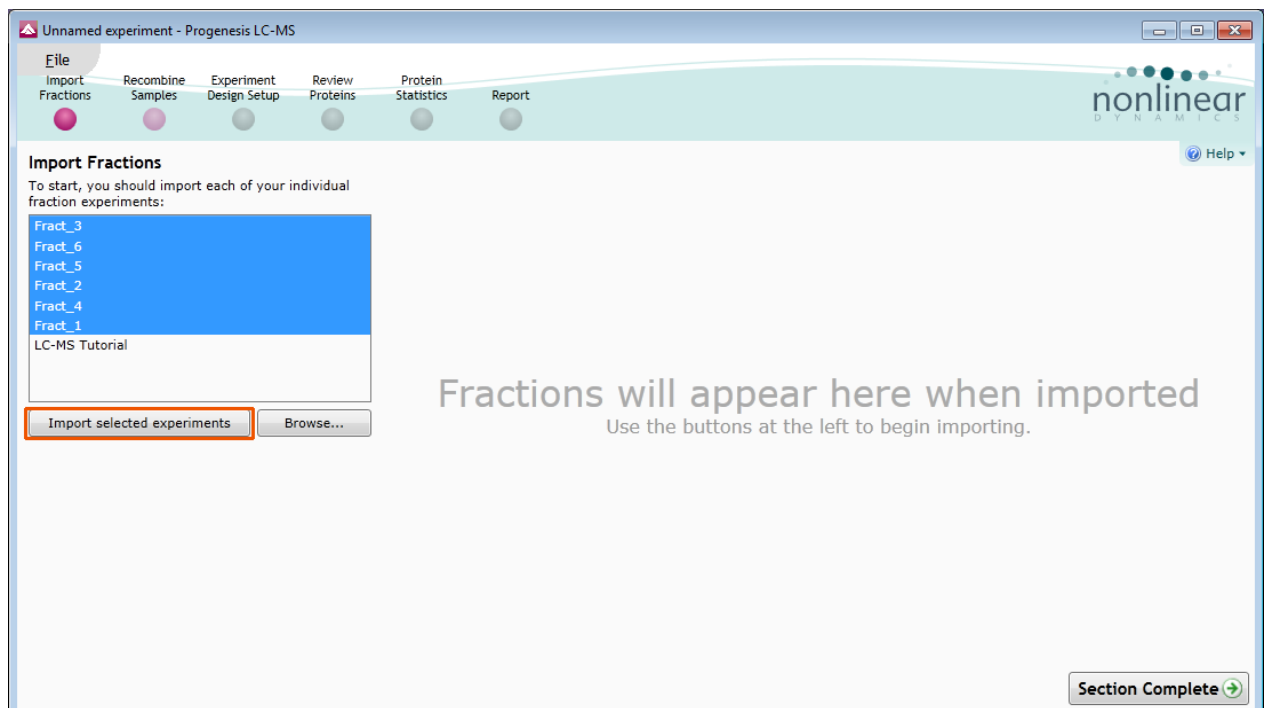
The combining of the single fraction experiments is performed in the second stage of the Fractionation workflow.

Select **Recombine analysed fractions...**



## Stage 1 Import Fractions

The Import Fractions stage of the work opens, select the experiments that correspond to the (6) fractions.



Click **Import selected experiments**

Details for each single fraction experiments appear in the panel showing numbers of identified peptides.

Unnamed experiment - Progenesis LC-MS

File Import Fractions Recombine Samples Experiment Design Setup Review Proteins Protein Statistics Report

**Import Fractions**  
To start, you should import each of your individual fraction experiments:

Fract\_3  
Fract\_6  
Fract\_5  
Fract\_2  
Fract\_4  
Fract\_1  
LC-MS Tutorial

Import selected experiments Browse...

**Peptides per fraction**

500  
400  
300  
200  
100  
0

1 2 3 4 5 6

Fraction

To obtain the correct peptide distribution, put the fractions in order:

1 **Fract\_3** ×  
368 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

2 **Fract\_6** ×  
146 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

3 **Fract\_5** ×  
320 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

4 **Fract\_2** ×  
318 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

5 **Fract\_4** ×  
454 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

6 **Fract\_1** ×  
20 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

Move Fract\_4: [down] [up] Help

Section Complete

You can adjust the order of the single fraction experiments to reflect the order of the fractions by dragging the single fraction experiments to the correct position.

Unnamed experiment - Progenesis LC-MS

File Import Fractions Recombine Samples Experiment Design Setup Review Proteins Protein Statistics Report

**Import Fractions**  
To start, you should import each of your individual fraction experiments:

Fract\_3  
Fract\_6  
Fract\_5  
Fract\_2  
Fract\_4  
Fract\_1  
LC-MS Tutorial

Import selected experiments Browse...

**Peptides per fraction**

500  
400  
300  
200  
100  
0

1 2 3 4 5 6

Fraction

To obtain the correct peptide distribution, put the fractions in order:

1 **Fract\_2** ×  
318 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

2 **Fract\_3** ×  
368 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

3 **Fract\_4** ×  
454 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

4 **Fract\_1** ×  
20 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

5 **Fract\_5** ×  
320 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

6 **Fract\_6** ×  
146 peptides identified in 6 runs  
Normalised: Yes  
Last saved: 04/08/2011

Move Fract\_1: [down] [up] Help

Section Complete

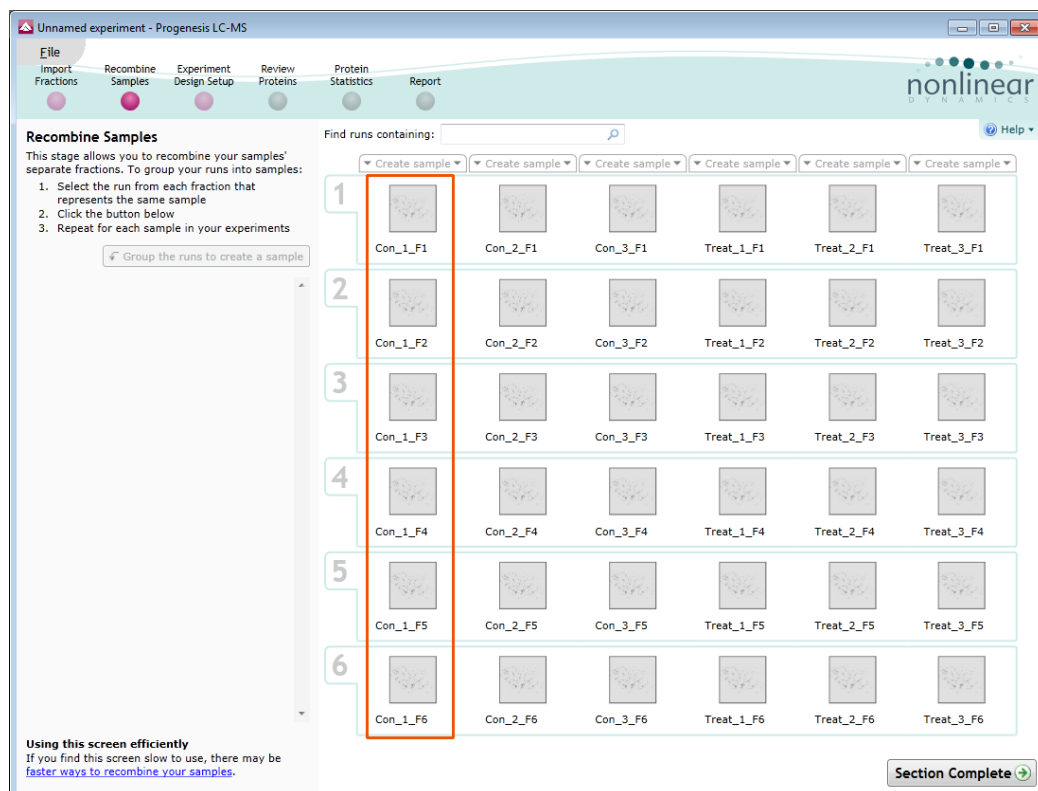
**Note:** the graph of 'peptides per fraction' updates to reflect the new order.

To move to the **Recombine Samples** stage click **Section Complete**.

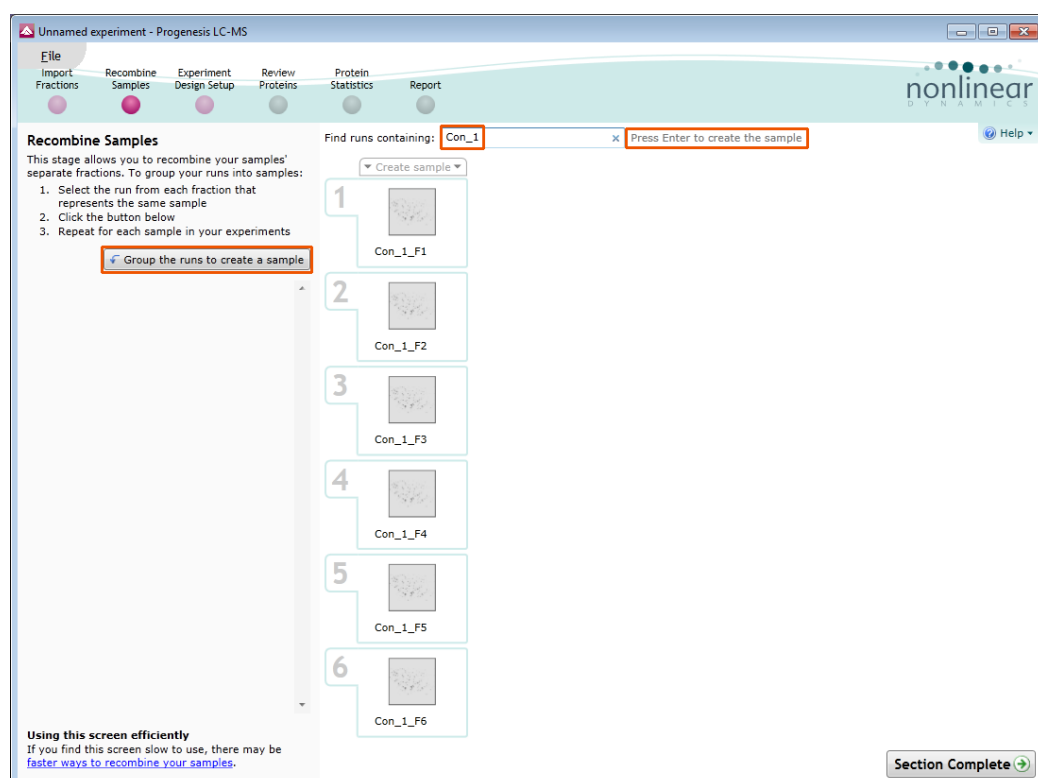
## Stage 2 Recombine Samples

At this stage you will recombine the samples by selecting the runs that correspond to each sample from the single fraction experiments.

**Note:** how efficiently you use this page will depend on how methodically you have named the various sample runs. For this example Samples are Con\_1, 2 and 3 and Treat\_1, 2 and 3

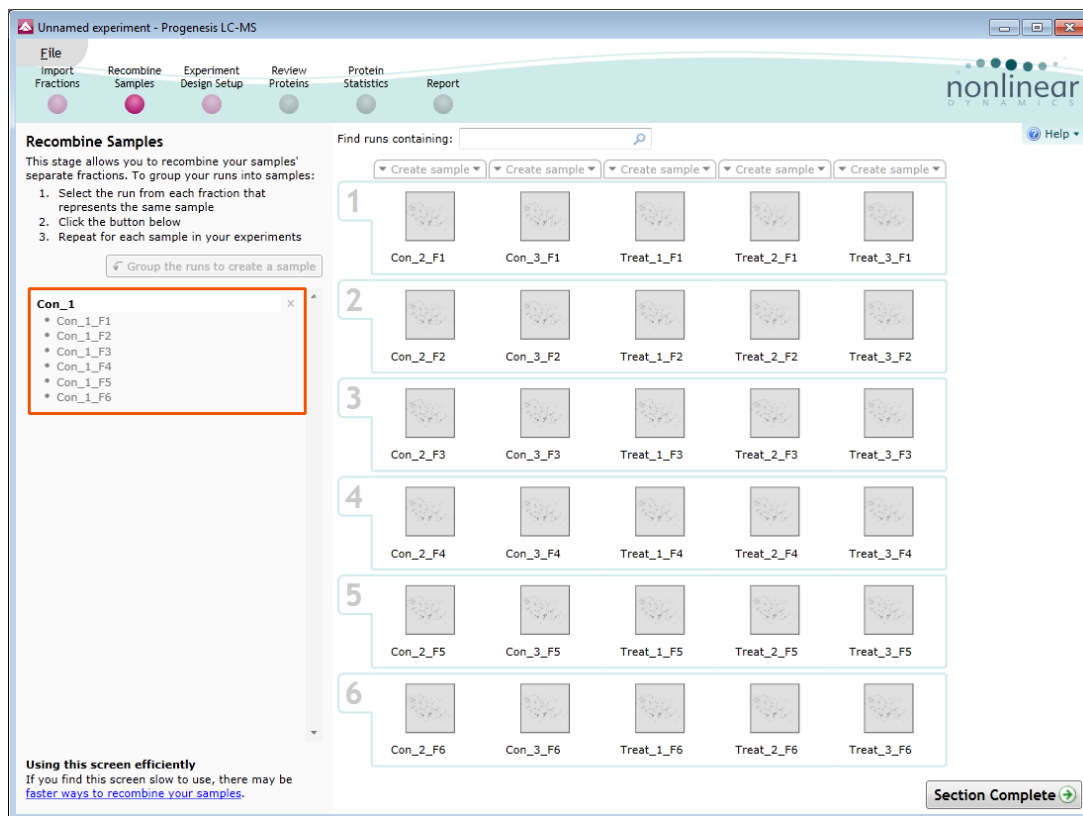


Typing **Con\_1** in the **Find runs containing** search box will locate the runs corresponding to sample Con\_1.

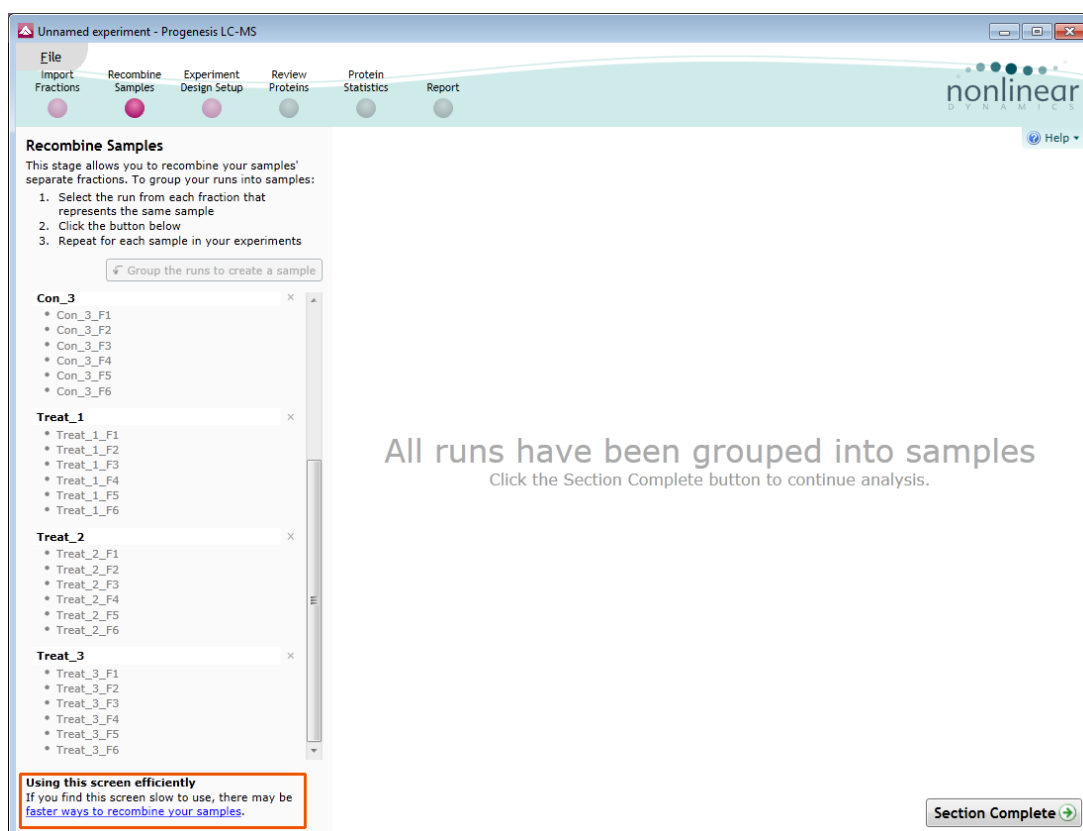


Click 'Enter' or **Group the runs to create a sample** in the left hand panel

**Note:** using Enter will set the sample name as **Con\_1**, overtyping to rename as required.



Repeat for the remaining samples.



**Note:** as mentioned before other ways of Recombining the samples can be applied, depending on the naming conventions used; use the link, bottom left, to see the alternatives.

Having completed the recombination of the samples, click **Section Complete**.

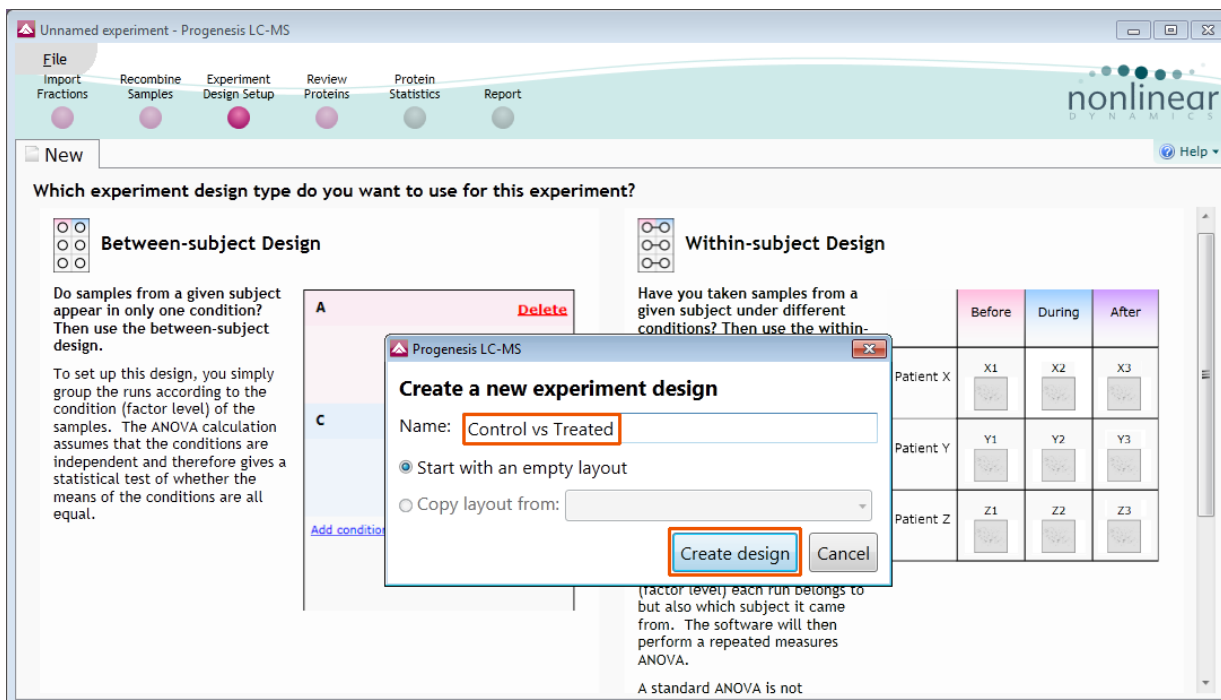
**Note:** at this point the data is re-normalised to account for the fractionation of the samples.

## Stage 3 Experiment Design Setup

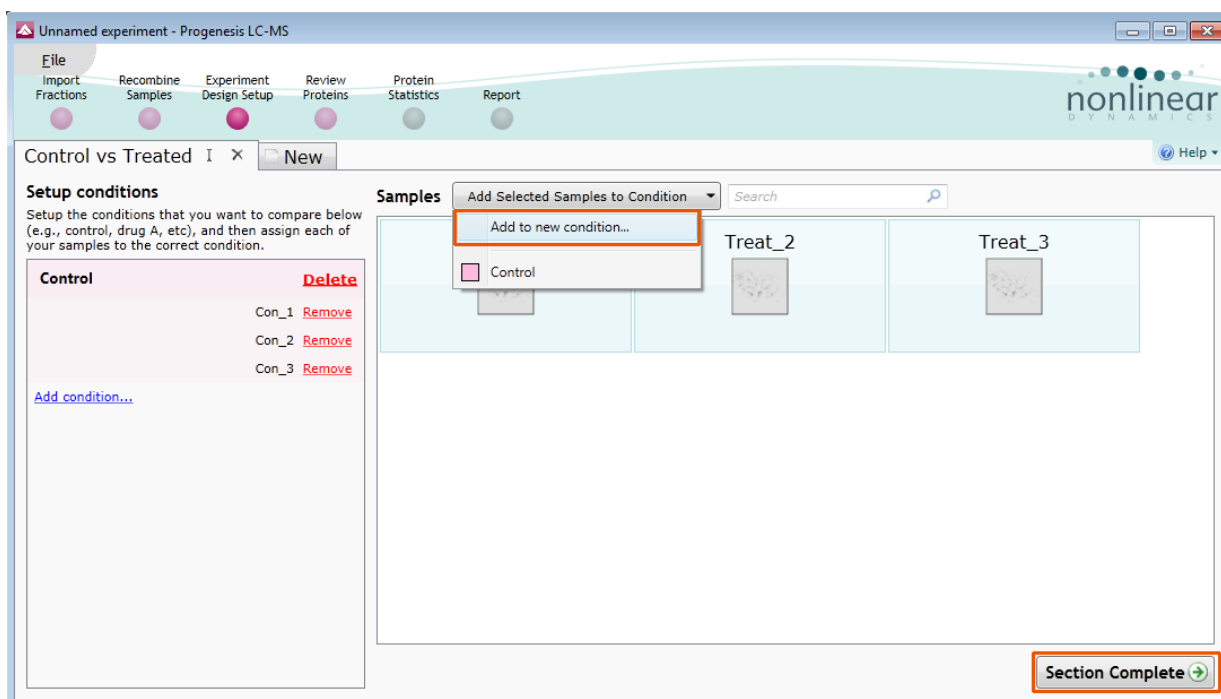
Having recombined your samples you can now define the experimental designs most appropriate for your original experiment.

This experiment contains 2 conditions: Control and Treated and uses the **Between-subject design** to group the analysed samples to reflect the Biological conditions in the original study.

Select **Between-subject** and give design an appropriate name.



Highlight the samples, to add them on to a new condition click on **Add Selected Samples to Condition**

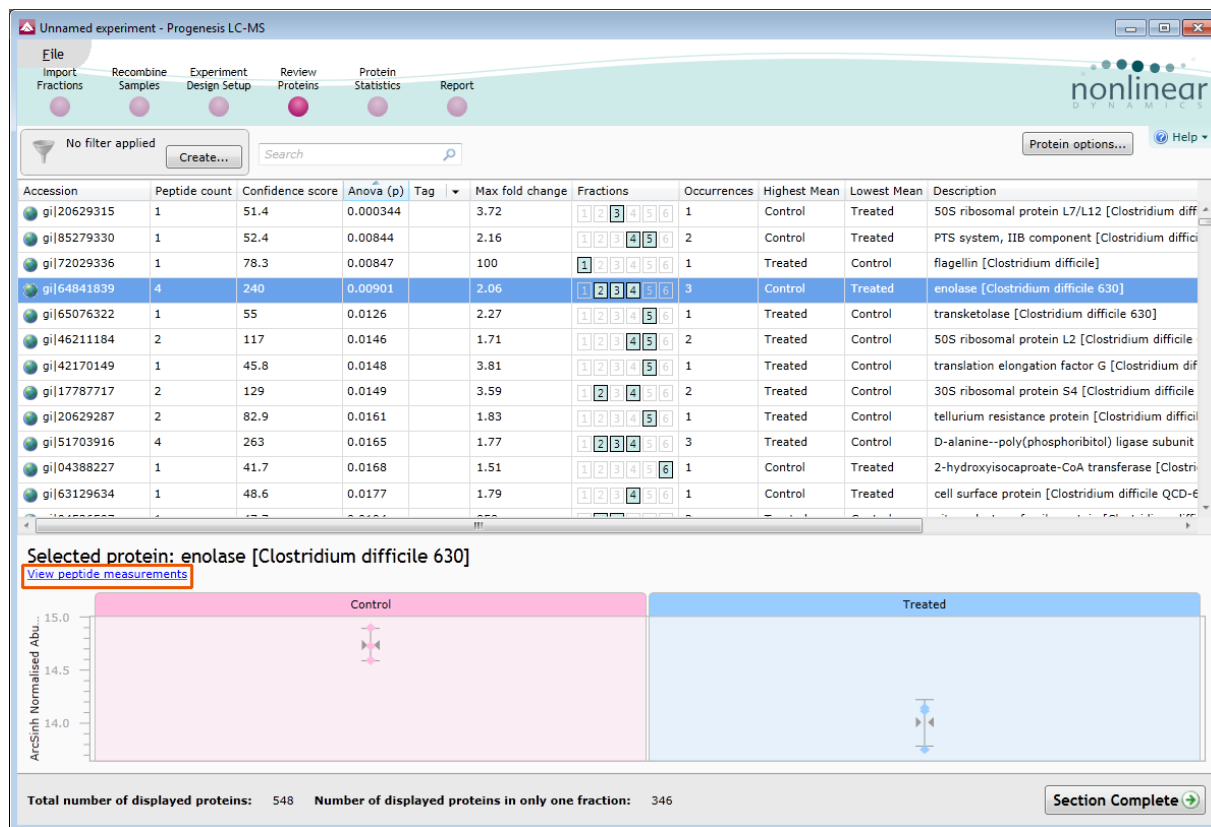


Click **Section Complete** to move to Protein View

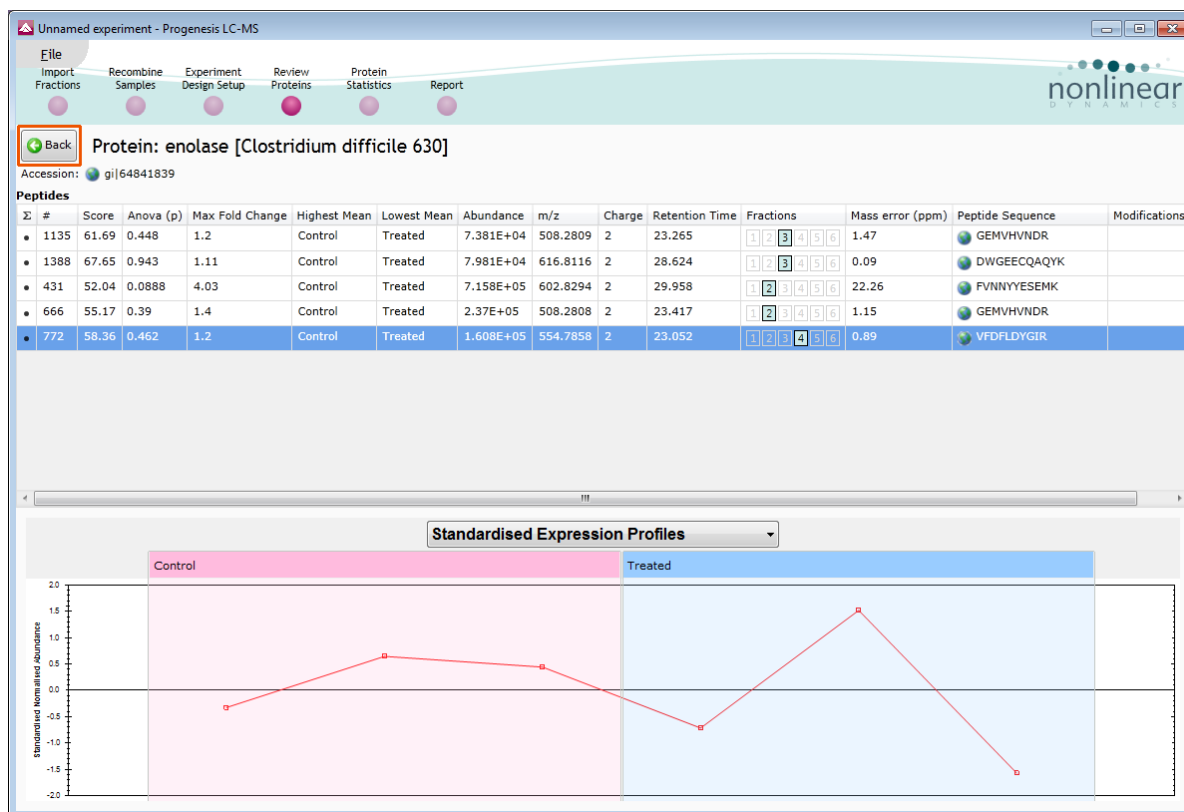


## Stage 4 Review Proteins

The recombined data can now be viewed at the level of the Proteins.



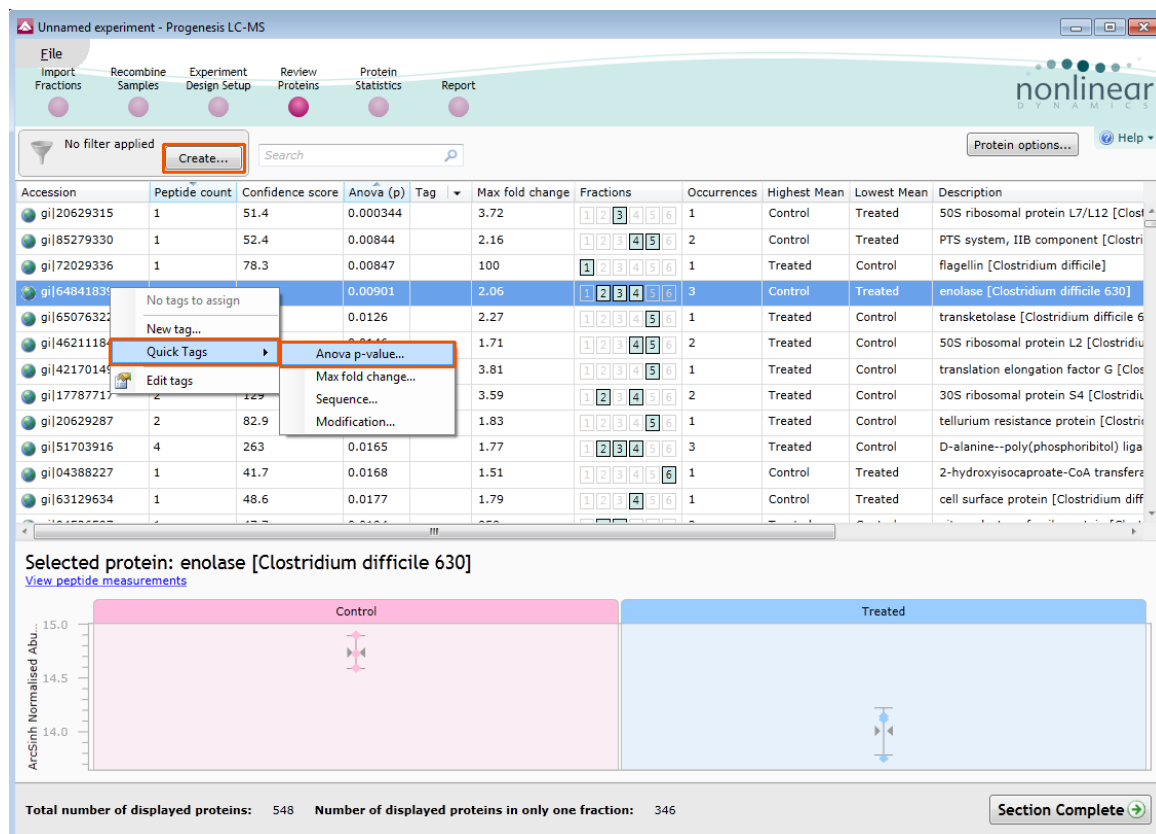
And at the peptide level when you click **view peptide measurements**



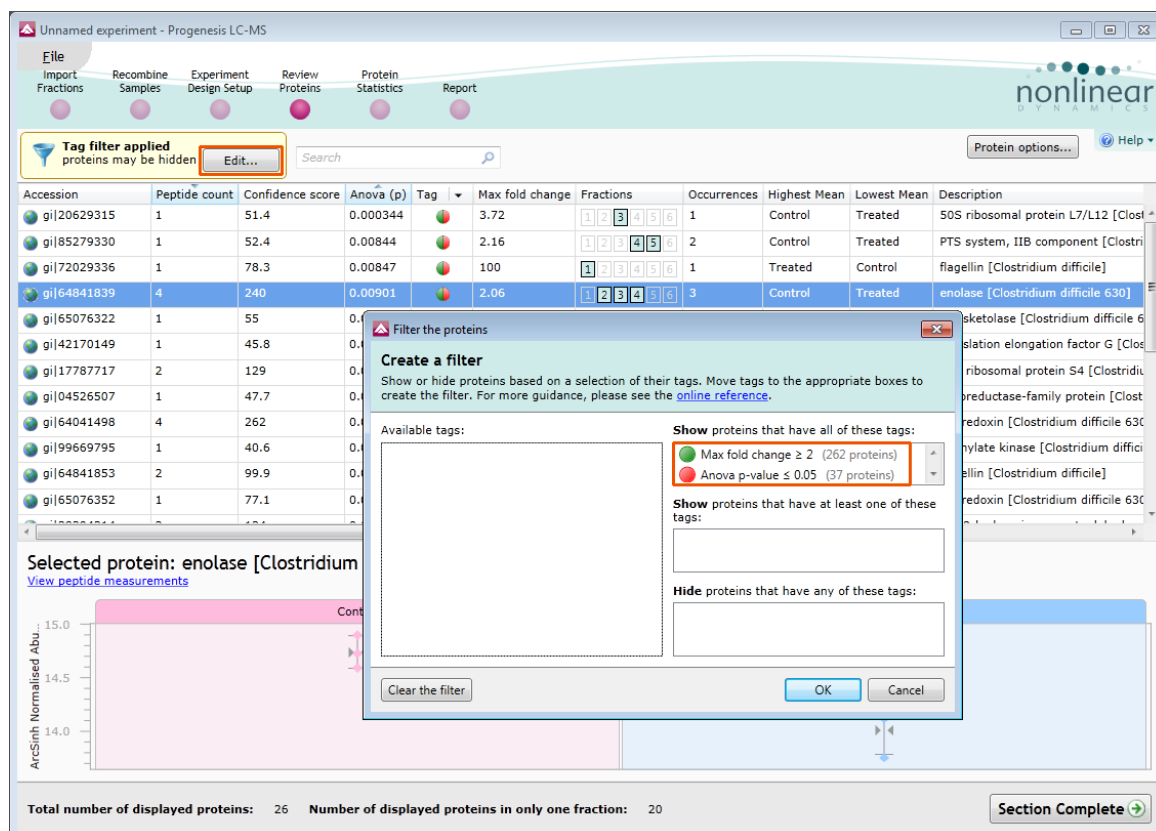
Click back to return to the **Protein View**

Using the Protein Tags you can generate a list of proteins based on similar properties and thresholds.

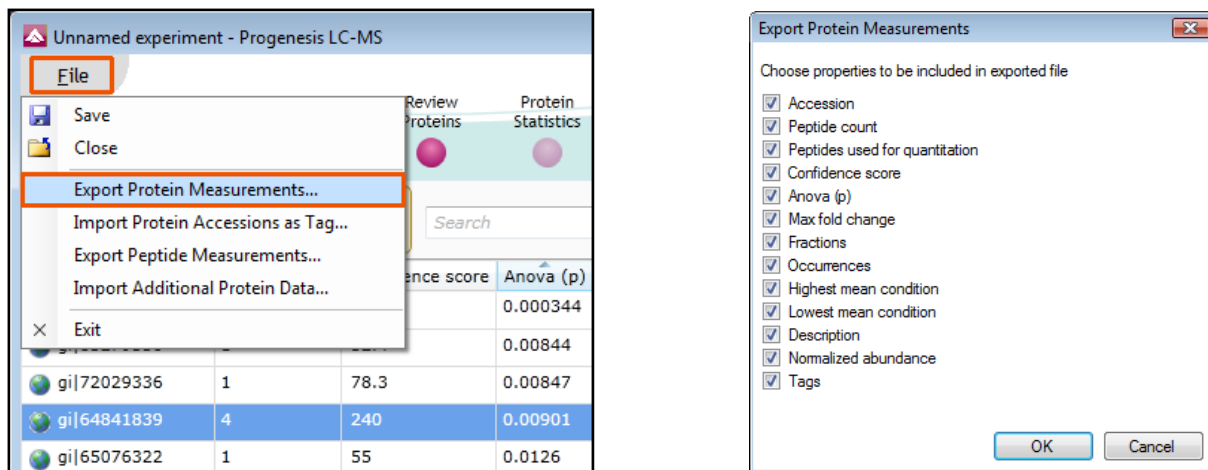
Right click on a protein in the table and use the Quick Tags to generate tags for proteins with **Anova p-value  $\leq 0.05$**  and a **Max fold change  $\geq 2$** .



Using the 'Tag' filters the list is reduced to the relevant proteins. Details of these proteins can be exported by selecting **Export Protein Measurements** from the **File** menu.



Details of these proteins can be exported by selecting **Export Protein Measurements** from the **File** menu



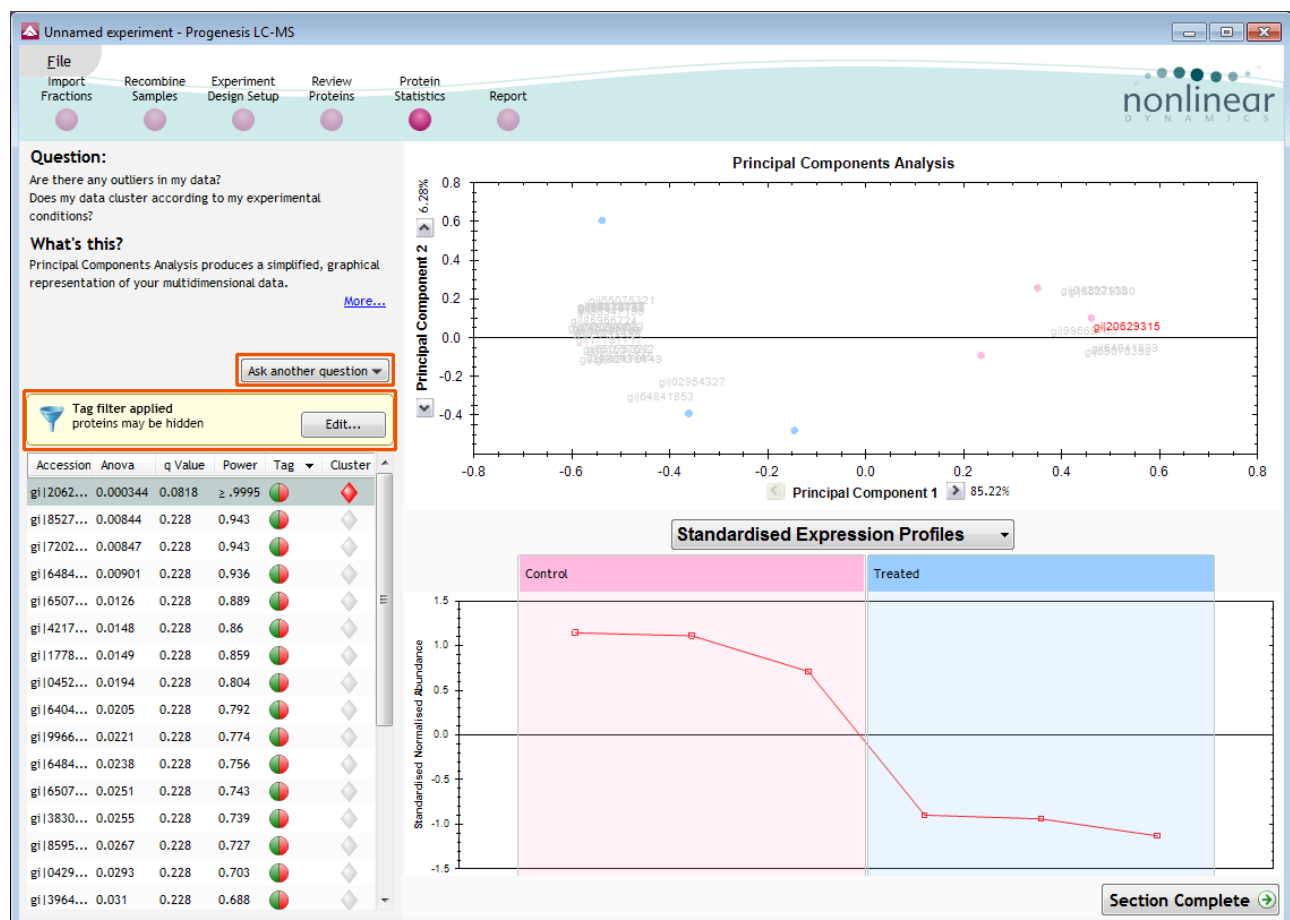
A dialog allows you to control the details of the output file.

Now move to the Protein Statistics section by clicking on **Protein Statistics** icon on the workflow at the top of the screen.

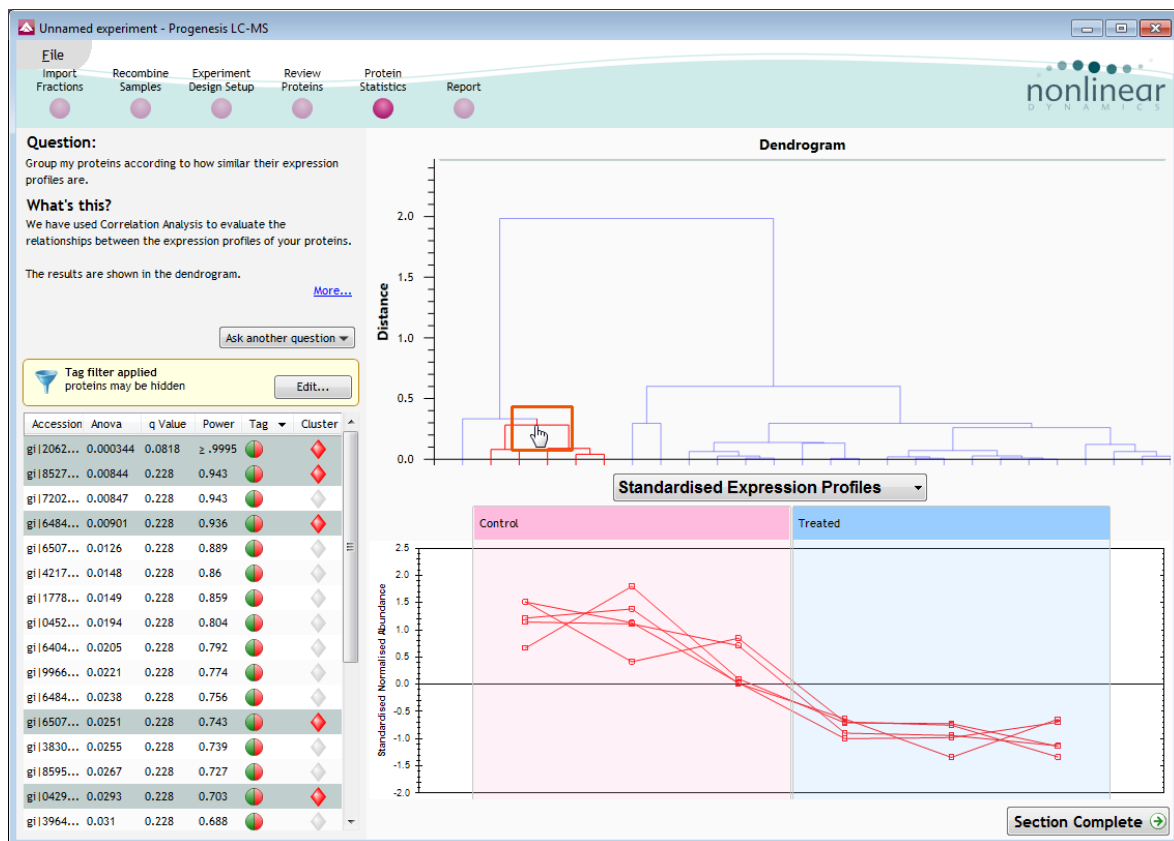
Finally your recombined data can be saved as a **Multi-fraction experiment**.

## Stage 5: Protein Statistics

Protein Statistics opens with a Principal Components Analysis (PCA) for all the proteins displayed.



The Multivariate Stats can now be applied to all or subsets of proteins as determined by the current Tag filters. Allowing you to identify similar patterns of expression using the Correlation Analysis.

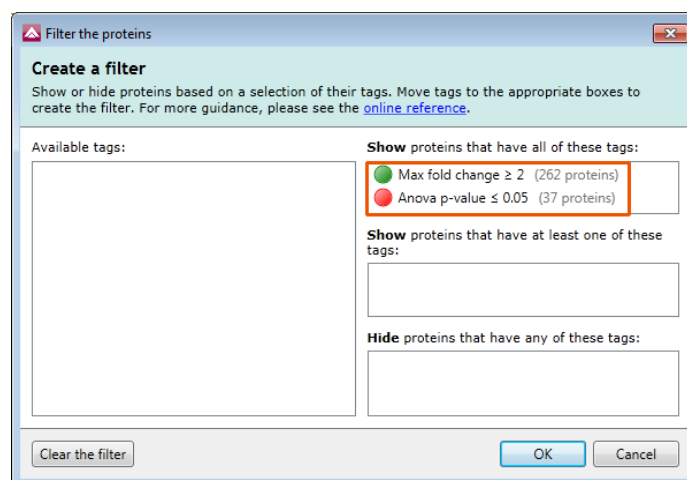


Now move to the **Report** section to report on Proteins and /or peptides.

## Stage 6: Reporting

The **Report Design** stage allows you to select what views you want to include in a report based on the list of **currently selected proteins**.

As an example we will create a report for **only** the proteins with **Anova p-value**  $\leq 0.05$  and a **Max fold change**  $\geq 2$ .



1. First reduce the proteins to report on by selecting the tags: **Anova p-value**  $\leq 0.05$  and a **Max fold change**  $\geq 2$ .
2. Expand the various Report Design options (by default they are all selected)
3. Un-tick as shown below

4. Click **Create Report**

**Proteins**

Tag filter applied  
features may be hidden

Accession	Anova (p)	Fold	Tag	Description
gi 20629315	0.000344	3.7		50S ribosomal protein L7/L12 [Clostridium difficile 630]
gi 85279330	0.00844	2.2		PTS system, IIB component [Clostridium difficile 630]
gi 72029336	0.00847	100.4		flagellin [Clostridium difficile]
gi 64841839	0.00901	2.1		enolase [Clostridium difficile 630]
gi 65076322	0.0126	2.3		transketolase [Clostridium difficile 630]
gi 42170149	0.0148	3.8		translation elongation factor G [Clostridium difficile 630]
gi 17787717	0.0149	3.6		30S ribosomal protein S4 [Clostridium difficile 630]
gi 04526507	0.0194	858.0		nitroreductase-family protein [Clostridium difficile 630]
gi 64041498	0.0205	3.3		thioredoxin [Clostridium difficile 630]
gi 99669795	0.0221	15.3		adenylate kinase [Clostridium difficile 630]
gi 64841853	0.0238	2.2		flagellin [Clostridium difficile]
gi 65076352	0.0251	4.8		thioredoxin [Clostridium difficile 630]
gi 38304314	0.0255	3.9		(R)-2-hydroxyisocaproate dehydrogenase [Clostridium difficile]
gi 85956724	0.0267	2.6		cell surface protein [Clostridium difficile 630]
gi 04292138	0.0293	2.5		flagellin [Clostridium difficile]
gi 39642159	0.031	8.3		pyruvate-flavodoxin oxidoreductase [Clostridium difficile QCD-23m63]
gi 48078745	0.0331	2.1		aspartate aminotransferase [Clostridium difficile 630]
gi 00347190	0.0359	2.6		cell surface protein [Clostridium difficile 630]
gi 82751345	0.0423	2.8		PTS system, IIB component [Clostridium difficile 630]
gi 02954327	0.0439	2.9		cell surface protein [Clostridium difficile 630]
gi 27875768	0.0445	3.4		ferredoxin-NADP(+) reductase subunit alpha [Clostridium difficile 630]
gi 21210623	0.046	3.4		transcription elongation factor [Clostridium difficile 630]

**Report Design**

Title: Significantly Changing Proteins

Select the sections you wish to include in your report:

- ☐ Experiment design
- ☒ Protein report
  - Include tables showing protein abundances and peptides identified for each protein
    - ☒ Protein table
    - ☒ Peptide tables
- ☒ Protein Details
  - Reports the full details of every protein which matches your current filter
    - ☒ Tags
    - ☒ Expression profile

**Create Report**

This opens a dialog to allow you to save the report, after which it will be opened in the form of a web page.

Significantly Changing Proteins									
<b>Experiment: Unnamed experiment</b>									
<b>Report created: 05/08/2011 15:17:14</b>									
Proteins									
Protein building options									
Protein grouping <b>Group similar proteins</b>									
Protein quantitation <b>Using only features with no protein conflicts</b>									
Accession	Peptides	Score	Anova (p)	Fold	Tags	Fractions	Description	Average Normalised Abundances	
								Control	Treated
gi 64041498	4	261.88	0.02	3.31		1 2 3 4 5 6	thioredoxin [Clostridium difficile 630]	2.34e+005	7.73e+005
gi 64841839	4	239.74	9.01e-003	2.06		1 2 3 4 5 6	enolase [Clostridium difficile 630]	1.27e+006	6.14e+005
gi 38304314	2	133.82	0.03	3.90		1 2 3 4 5 6	(R)-2-hydroxyisocaproate dehydrogenase [Clostridium difficile]	5.53e+004	2.16e+005
gi 17787717	2	128.88	0.01	3.59		1 2 3 4 5 6	30S ribosomal protein S4 [Clostridium difficile 630]	8.94e+004	3.21e+005
gi 64841853	2	99.94	0.02	2.22		1 2 3 4 5 6	flagellin [Clostridium difficile]	1.16e+005	2.58e+005
gi 05386167	2	95.22	0.05	2.54		1 2 3 4 5 6	small acid-soluble spore protein A [Clostridium difficile 630]	9.35e+004	2.37e+005
gi 48078745	2	94.59	0.03	2.11		1 2 3 4 5 6	aspartate aminotransferase [Clostridium difficile 630]	2.71e+005	5.72e+005
gi 85956724	1	82.04	0.03	2.59		1 2 3 4 5 6	cell surface protein [Clostridium difficile 630]	1.20e+005	3.11e+005
gi 72029336	1	78.27	8.47e-003	100.38		1 2 3 4 5 6	flagellin [Clostridium difficile]	2040.73	2.05e+005

Click on the **Accession No.** in the proteins section of the Report and this will take you to the Assigned peptides for this protein

gi | 64841839

enolase [Clostridium difficile 630]  
4 peptides

Sequence	Feature	Score	Hits	Mass	Charge	Fraction	Modifications	In quantitation	Average Normalised Abundances	
									Control	Treated
DWGEECAQYK	1388	67.65	2	1231.6086	2	1 2 3 4 5 6		yes	7.98e+004	7.18e+004
FVNNYSEMK	431	52.04	1	1203.6443	2	1 2 3 4 5 6		yes	7.16e+005	1.78e+005
GEMVHVNDNR	666	55.17	5	1014.5470	2	1 2 3 4 5 6		yes	2.37e+005	1.70e+005
GEMVHVNDNR	1135	61.69	2	1014.5473	2	1 2 3 4 5 6		yes	7.38e+004	6.17e+004
VDFLDYGIR	772	58.36	1	1107.5570	2	1 2 3 4 5 6		yes	1.61e+005	1.33e+005

gi | 38304314

(R)-2-hydroxyisocaproate dehydrogenase [Clostridium difficile]  
2 peptides

Sequence	Feature	Score	Hits	Mass	Charge	Fraction	Modifications	In quantitation	Average Normalised Abundances	
									Control	Treated
AFSCMSEGLSCK	1747	78.62	7	1279.6782	2	1 2 3 4 5 6		yes	4.74e+004	1.65e+005
FCPWWCHYTHER	4756	55.20	1	1425.6537	2	1 2 3 4 5 6		yes	7844.75	5.06e+004

Having closed the report it can be reopened by double clicking on the saved html file.

**Note:** you can also copy and paste all or selected sections of the report to Excel and/or Word.

**Note:** there are separate panels for reporting on Proteins and Peptides

Finally your recombined data can be saved as a **Multi-fraction experiment**.

## Stage 7: Saving a Multi-Fraction experiment

When you opened the Fractionation workflow and started working with the recombination of your 'Single Fraction Experiments' the workflow recognises the current experiment as '**Unnamed**' this status will change as you close and/or save the experiment.

On saving the new multi-fraction experiment appears on the **Combine analysed fractions** page

**Progenesis LC-MS**

File

Experiments

Perform analysis | Combine analysed fractions

**Fractionation**

If you're using fractionation to [increase peptide coverage](#), this workflow will help you to recombine your fractionated samples' data.

For more information on how Progenesis LC-MS analyses fractionated samples, [visit the website](#).

- Analyse each fraction separately**  
Create a separate experiment for each fraction. Each experiment will contain a single run for each of your samples.  
[Analyse a single fraction...](#)
- Recombine the fractionated samples**  
After analysing each fraction separately, these can be added to a multi-fraction experiment to recombine your samples.  
[Recombine analysed fractions...](#)

[View online tutorial](#)

**Recent multi-fraction experiments**

Search

**My Recombination**  
In folder: D:\Customer Data\LC-MS\Fractionated Experiments  
Has 6 fractions, with 571 proteins identified across 6 samples.  
[Open](#) [Options](#)

**Getting started with Progenesis**  
Here are some resources to help you get started with Progenesis LC-MS:

- [Data analysis with Progenesis LC-MS](#)
- [Tutorial and user guide](#)
- [Frequently-asked questions](#)

**Double the proteome coverage?**  
Gas phase fractionation can vastly increase protein and proteome coverage.

**Optimise Gas Phase Fractionation**  
If this experiment is a pilot run for [gas phase](#) number of fractions to see the optimal m/z

Number of fractions: 5

Fraction	Start m/z	End m/z
1	440	593
2	593	688
3	688	804

[Browse for a multi-fraction experiment...](#)

Details of the new experiments location (folder), number of fractions, samples and total identified proteins.

**Note:** you can reopen the experiment by either double clicking on it or using **open**.

This completes a guided tour through using Progenesis LC-MS to analyse fractionated data.

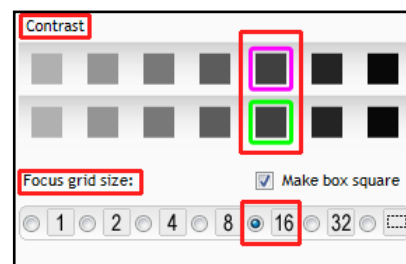


## Appendix 1: Manual assistance of Alignment

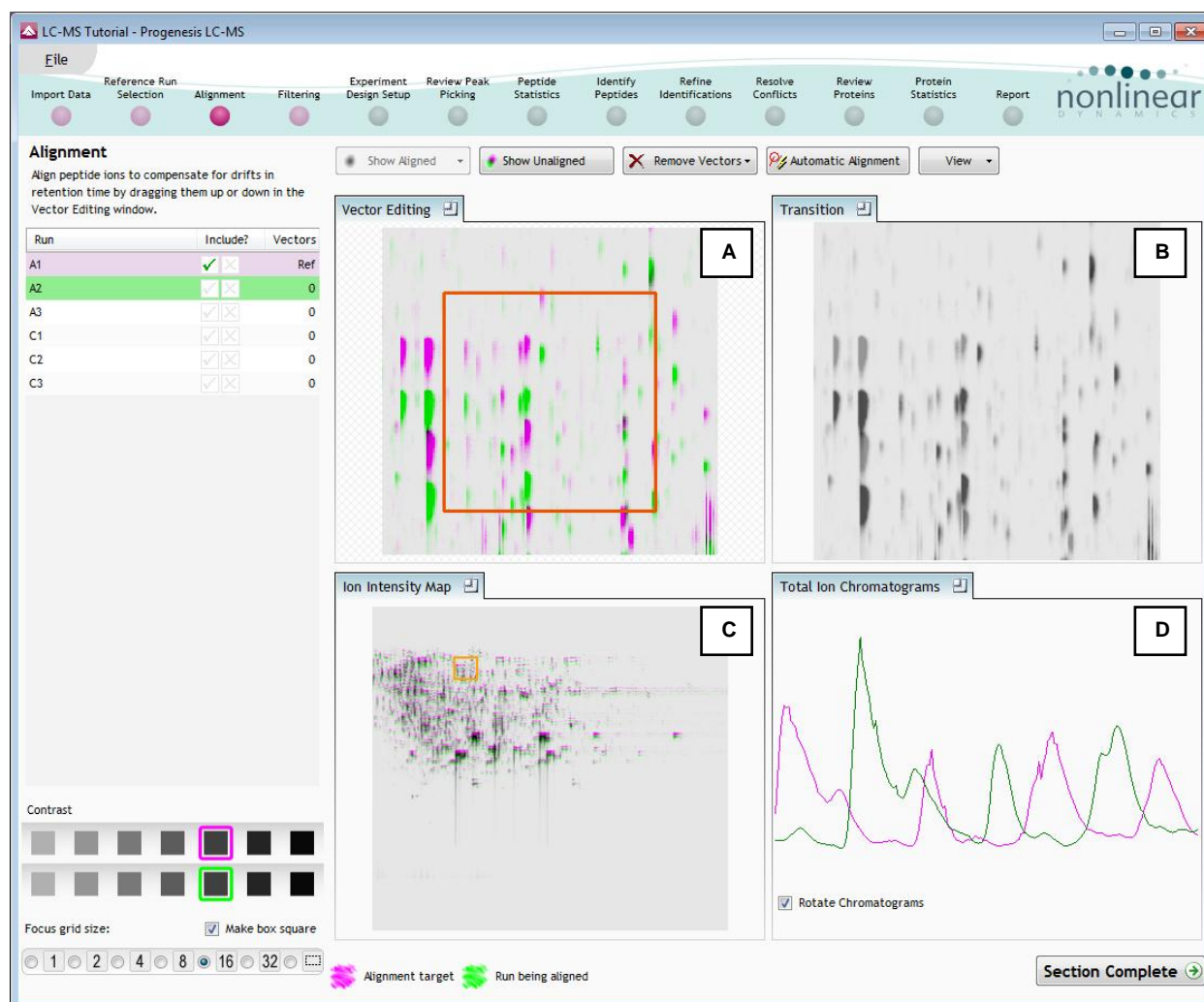
### Approach to alignment

To place manual alignment vectors on a run (A2 in this example):

1. Click on Run A2 in the **Runs** panel, this will be highlighted in green and the reference run (A1) will be highlighted in magenta.
2. You will need approximately 5 - 10 **alignment vectors** evenly distributed from top to bottom of the whole run.
3. First ensure that the size of the focus area is set to **8** or **16** in the Focus grid size on the bottom left of the screen.



Click on an area (see below) in the **Whole Run** view (C) to refocus all the windows. Adjust Contrast as required.

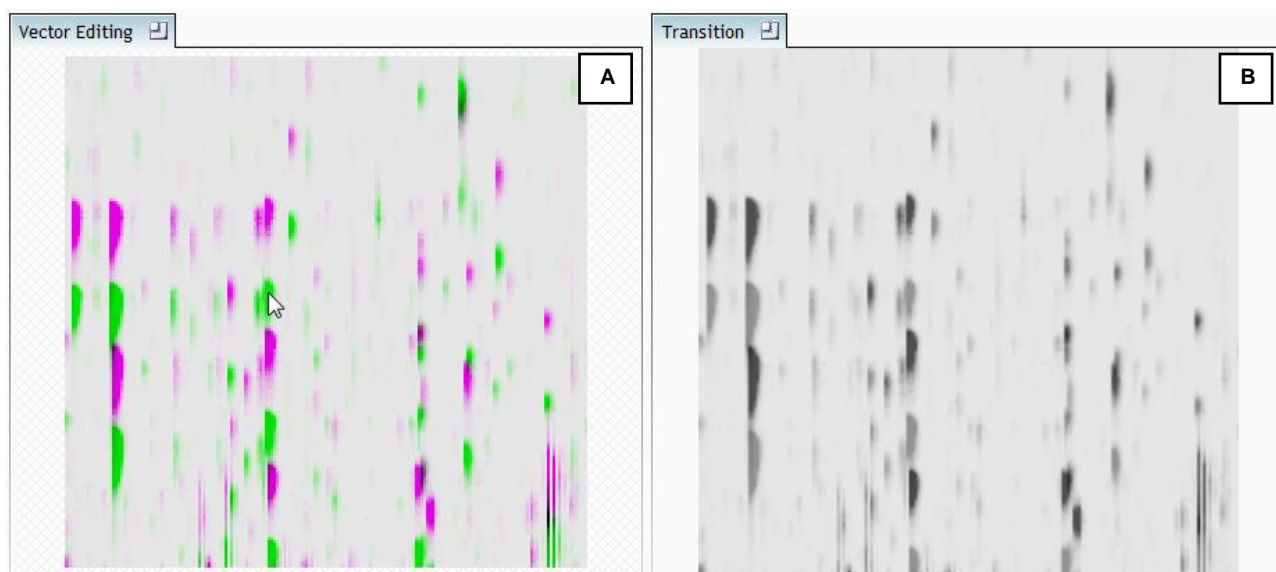


**Note:** the features moving back and forwards between the 2 runs in the **Transition** view indicating the misalignment of the two LC-MS runs

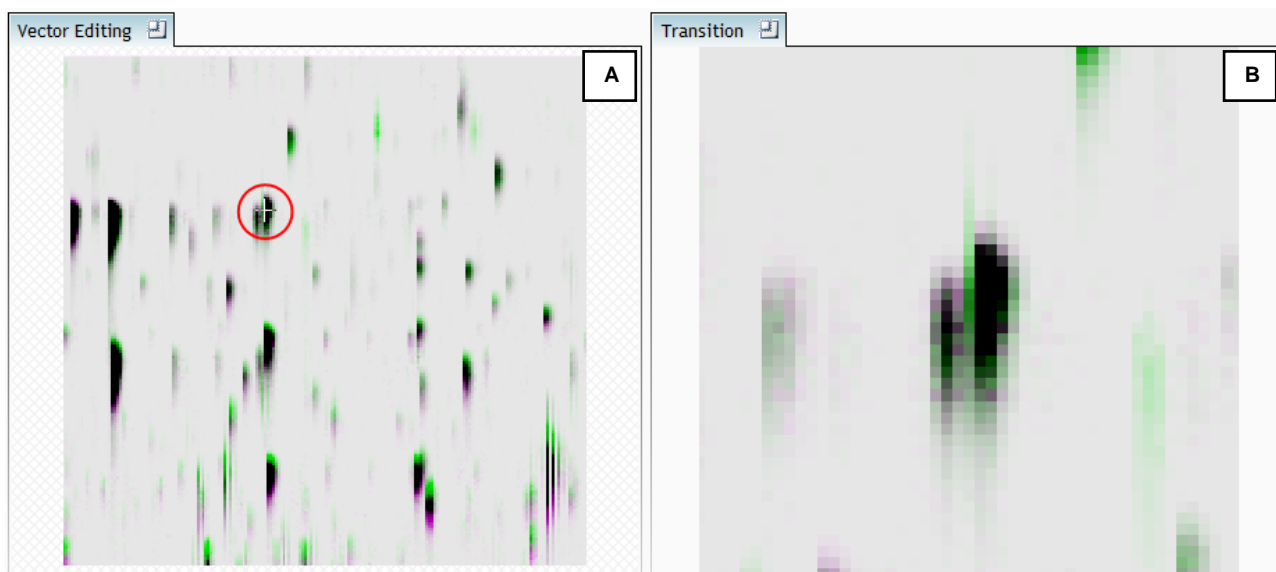
**Note:** The **Total Ion Chromatogram** view also reflects the misalignment of the 2 runs for the current Retention Time range (vertical dimension of the current Focus grid in the **Whole Run** view).



4. Click and hold on a green feature in Window A as shown below.

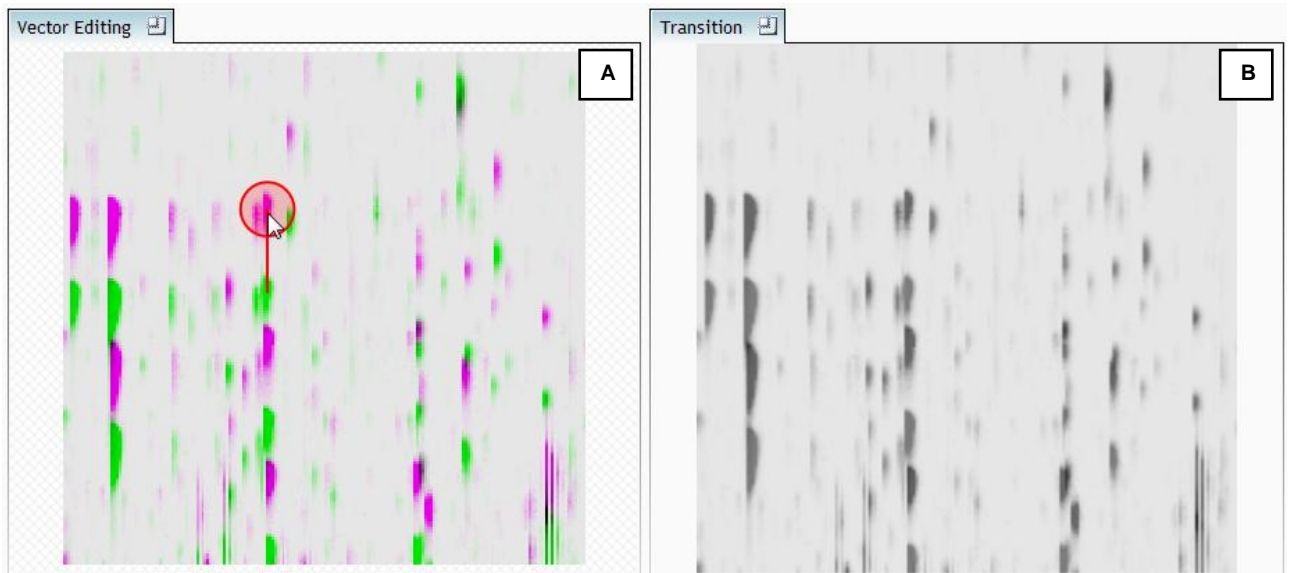


5. As you are holding down the left mouse button drag the green feature over the corresponding magenta feature of the reference run. The red circle will appear as shown below indicating that a positional lock has been found for the overlapping features.



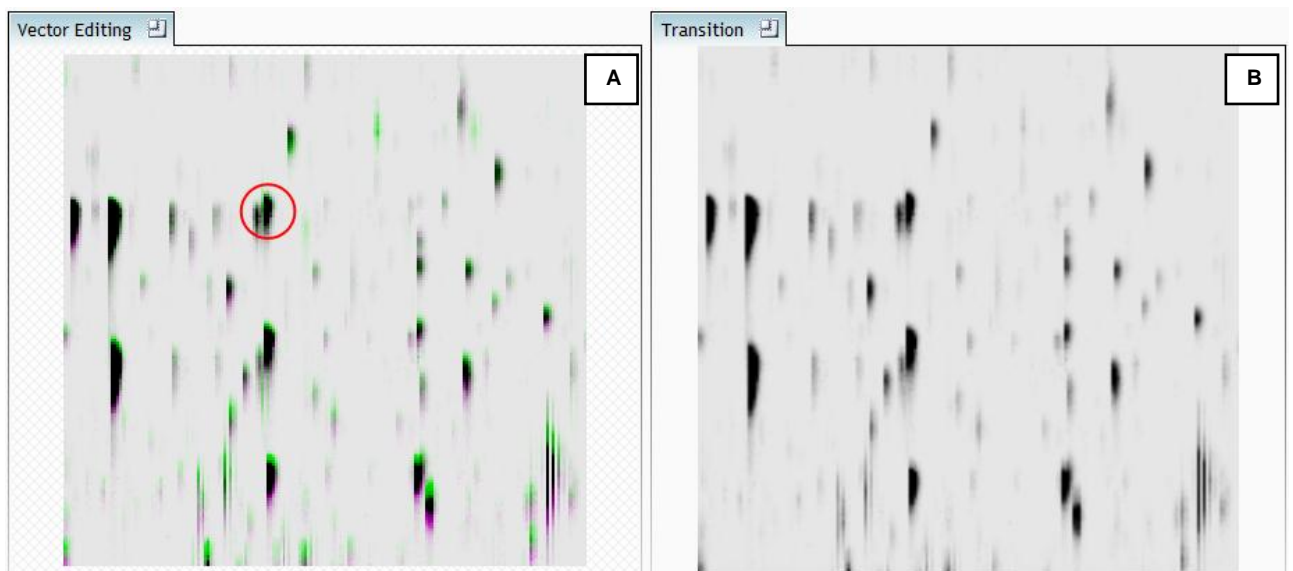
**Note:** as you hold down the mouse button, window B zooms in to help with the alignment.

6. On releasing the left mouse button the view will 'bounce' back and a red vector, starting in the green feature and finishing in the magenta feature will appear.

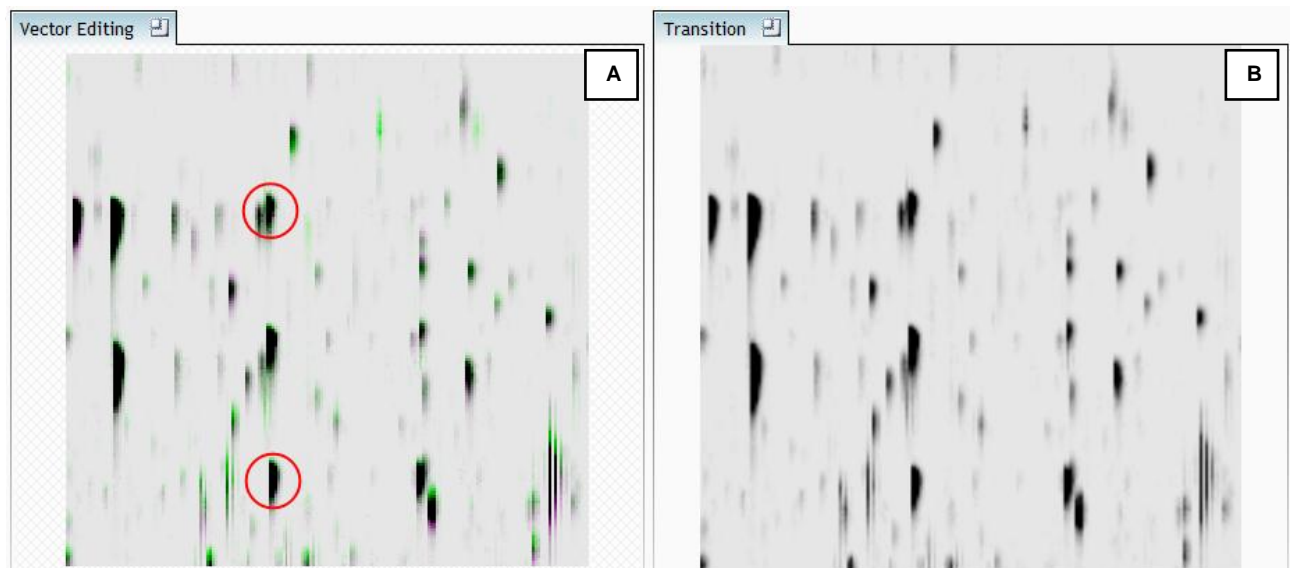


**Note:** an incorrectly placed vector is removed by right clicking on it in the **Vector Editing** window

7. Now click **Show Aligned** on the top tool bar to see the effect of adding a single vector.



8. Adding an additional vector will improve the alignment further. **Note** this time as you click to add the vector it 'jumps' automatically to the correct position using the information from the existing alignment vector.



Repeat this process moving the focus from top to bottom on the **Whole Run** view

**Alignment**

Align peptide ions to compensate for drifts in retention time by dragging them up or down in the Vector Editing window.

Run	Include?	Vectors
A1	<input checked="" type="checkbox"/>	Ref
A2	<input checked="" type="checkbox"/>	6
A3	<input checked="" type="checkbox"/>	0
C1	<input checked="" type="checkbox"/>	0
C2	<input checked="" type="checkbox"/>	0
C3	<input checked="" type="checkbox"/>	0

Contrast: ☐ ☐ ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐

Focus grid size: ☐ 1 ☐ 2 ☐ 4 ☐ 8 ☒ 16 ☐ 32 ☐ Make box square

Alignment target (pink) Run being aligned (green)

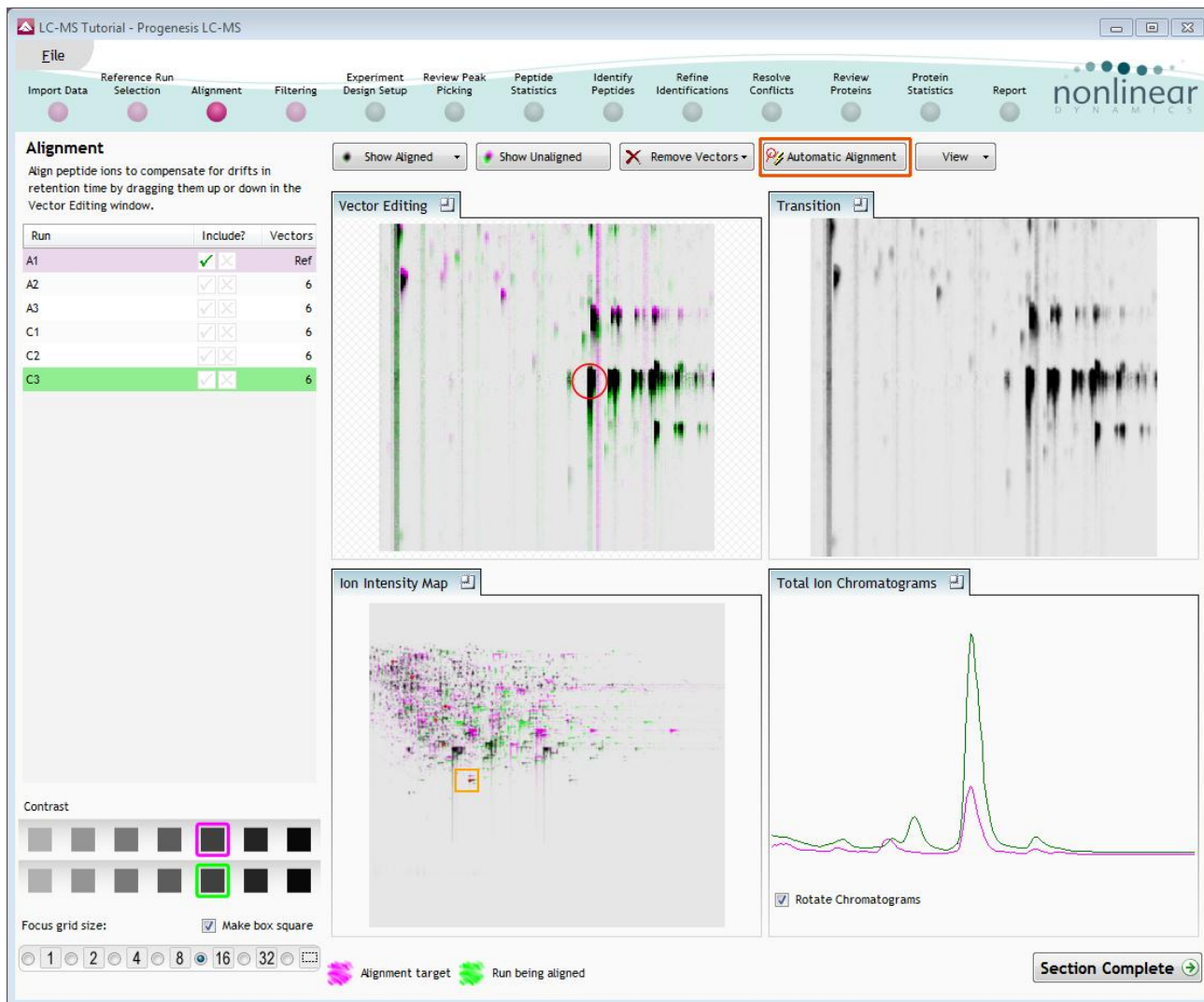
Section Complete

**Note:** the number of vectors you add is recorded in the **Runs** table

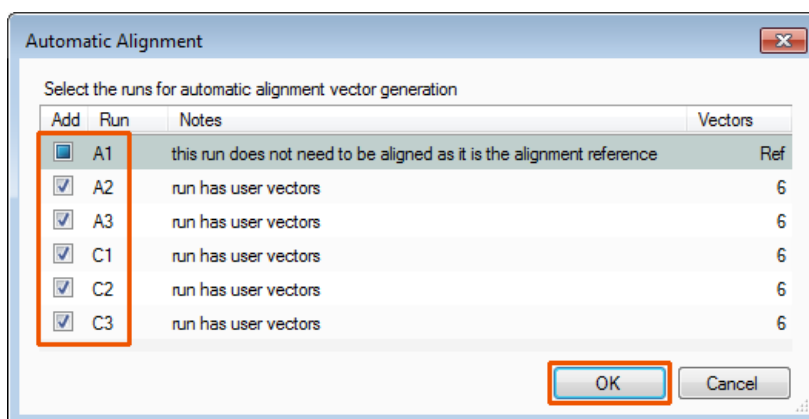
10. Repeat this process for all the runs to be aligned.

The number of manual vectors that you add at this stage is dependant on the misalignment between the current run and the Reference run. In many cases only using the Automatic vector wizard will achieve the alignment.

Also the 'ease' of addition of vectors is dependant on the actual differences between the LC-MS runs being aligned



11. Then select Automatic Alignment to bring up the Automatic Alignment dialog and click **OK**. The automatic alignment process will begin, using the manual vectors you have added to aid in automatic vector placement.



**Note:** the tick boxes next to the 'Run name control' which control whether vectors will be generated for each run.

To review the vectors, automatic and manual return to page 9

## Appendix 2: Within-subject Design

To create a **Within-subject Design** for your data set select this option on the **Experiment Design Setup** page and enter the name of the design.

In this example there are 3 Subjects (i.e. patients A, B and C) who have been individually sampled: Before(1), During (2) and After (3) treatment

LC-MS Tutorial - Progenesis LC-MS

File Reference Run Selection Alignment Filtering Experiment Design Setup Review Peak Picking Peptide Statistics Identify Peptides Refine Identifications Resolve Conflicts Review Proteins Protein Statistics Report

New Help

Which experiment design type do you want to use for this experiment?

☐ Between-subject Design

Do samples from a given subject appear in only one condition? Then use the between-subject design.

To set up this design, you simply group the runs according to the condition (factor level) of the samples. The ANOVA calculation assumes that the conditions are independent and therefore gives a statistical test of whether the means of the conditions are all equal.

**Within-subject Design**

Have you taken samples from a given subject under different conditions? Then use the within-subject design.

Note: you must have a sample from every subject for every condition to use a within-subject design.

**Create a new experiment design**

Name: Before During and After Treatment

☒ Start with an empty layout

☐ Copy layout from: [dropdown]

Create design Cancel

	Before	During	After
Patient X	X1	X2	X3
Patient Y	Y1	Y2	Y3
Patient Z	Z1	Z2	Z3

When the design page opens use the **Add Subject** and **Add Condition** buttons to create the matrix that fits your experimental design, over typing the names as required.



Then Drag and drop the Samples on to the correct 'cell' of the matrix.

**LC-MS Tutorial - Progenesis LC-MS**

File | Import Data | Reference Run Selection | Alignment | Filtering | Experiment Design Setup | Review Peak Picking | Peptide Statistics | Identify Peptides | Refine Identifications | Resolve Conflicts | Review Proteins | Protein Statistics | Report | nonlinear DYNAMICS

Before During and After Treatment | New | Help

**Setup conditions and subjects**

Setup the conditions and subjects for your experiment design on the right, and then assign each of your samples to the correct subject/condition cell in the grid.

1. Add a column for each condition.
2. Add a row for each subject.
3. Drag each of your samples to the correct location in the grid.

Filter samples:

	Before	During	After	Add Condition
Patient A	A1 	A2 	A3 	
Pateint B	B1 	B2 	B3 	
Patient C	C1 	<a href="#">Select Sample</a>	<a href="#">Select Sample</a>	

Add Subject

Section Complete

You can create additional Experimental Designs using the New tab

All of these Experimental Designs are available at all the following stages in the LC-MS workflow.